

Why did Challenger matter?

The loss last week of the US shuttle Challenger has in some sense shocked the world. It would be good if it helped the United States to a better programme.

THE anguish in the United States and around the world about last week's shuttle accident requires explanation. Why was this accident so much more compelling of public grief than, say, the loss of more than 200 US servicemen in an aircraft accident in Newfoundland just before Christmas or the other even more horrific accidents that studded 1985? There are two threads in any account of how last week's tragedy gripped the sombre imagination of the whole world, of which the first is the proof that Challenger provided that Marshall McLuhan's "global village" is now indeed a reality. The television cameras that recorded both the destruction of the spacecraft and the personal grief of the relatives of its occupants made personal tragedies general. The same television cameras brought the Vietnam war into the living rooms of the United States just over a decade ago, with awesome consequences. Last year, there was the famine in Ethiopia. Nobody will or should complain at the discovery, now, that the global village will increasingly haunt us.

The second reason for the sense of national tragedy in the United States last week is that Challenger carried with it into the Atlantic some of the sense of hope that the United States has cherished for its space programme. The competition in technology from Japan may have turned from an irritation to a cause for dismay, international problems that once seemed small, as in Central America, the Middle East and Central Europe, have turned out to be devilishly complicated, but there seemed, until last week, a high chance that the United States would win a new set of spurs for having made space travel by people into a routine happening. That there might nevertheless be an accident was always on the cards; reason says that the only question to answer was not "whether" but "when". But it is natural, and understandable, that people's pride should have given way to profound disappointment. Mourning is an appropriate mechanism of catharsis. That there has been, with the State of the Union message postponed a week, and the government's tight timetable with it. So what will happen now?

Future

President Reagan says the shuttle programme will continue, and his decision has been echoed, mostly with enthusiasm, throughout the United States. That, no doubt, is what the seven brave men and women who died with Challenger would have wished. Any other decision would turn the present sense of disappointment into one of outright failure. And, for strictly practical reasons, there is no choice. The Pentagon has been anxious about the reliability of the shuttle, and its turn-round time, for years, and is refurbishing some Titan rockets as satellite launchers until a batch of ten specialized rockets are delivered in 1988 or so; but, for some time, the three surviving models of the shuttle have a crucial job to do. The queue of commercial satellites waiting to be put in orbit is also long, and lengthening; and delay means more than mere cost. But this year was also intended to be a good year for launching scientific satellites, of which the Hubble Space Telescope (originally due to go up in August) is by far the most important. The scientific community, of which an influential part is fond of saying that the objectives of the space programme could all more surely and safely be

attained with unmanned rockets, should remember that this important instrument will require repeated visits by a shuttle crew during the fifteen years of its intended life.

So what lessons are to be learned? The hardest is that there are no simple lessons. No doubt the National Aeronautics and Space Administration will soon find the cause of last week's accident, modify the remaining spacecraft accordingly and restart the programme. In its short time at work, the shuttle has in any case amply demonstrated its usefulness, if not its effectiveness in economic terms. But one of the lasting casualties of last week's accident will be the administration's ambition to put the operation of the shuttle on a commercial footing. Another will be the notion that the entire ambition of the United States in space can be sustained by one chief launching instrument. It would be safer and probably also cheaper to keep the shuttle for the jobs that only the shuttle can carry out, using some other rocket to put everyday satellites into orbits. The obvious difficulty, in the year when the Gramm-Rudman Act will be playing havoc with the federal budget, is that there is no obvious place from which the extra funds can come. But it will be a great misfortune if last week's mood of resolution is too quickly lost in squabbling about budgets.

Plans

That is why it is sensible that the White House is taking a wider look at the future of the space programme as a whole. There are too many anomalies for comfort built into plans for the immediate future. Even Spacelab is an embarrassment, with its complicated arrangements for anticipating what the future could hold by running electrophoresis experiments whose time has not yet come. But the United States is more ambitiously committed to the concept of building a space station somewhere up there. Would it not make sense to let that plan lapse? Certainly the lobby that has been building up to send some kind of mission to Mars had better go into hiding for a while. None of this implies that the successes of the past few years now count for nothing; it is merely that they have enjoyed a run of luck, statistical in nature, that could not have been anticipated. To succeed with the help of luck is no shame, of course, but to expect that the luck will never run out is unwise.

There was particular sadness last week about the death of Christa McAuliffe, the New Hampshire schoolteacher on the shuttle. She had responded to President Reagan's offer, during the 1984 election campaign, to send a schoolteacher into space as an inspiration to students and, perhaps, also to the National Education Association, which seemed to be on the point of endorsing Mr Walter Mondale, Mr Reagan's chief opponent for the presidency. The unpalatable truth is that the venture was a publicity stunt of a kind too often foisted on the US space programme in the past, sometimes by its own managers. What last week's tragedy shows is that there is now no case for carrying on shuttle flights passengers whose presence is not essential. There should be no more schoolteachers, congressmen or even journalists (see *Nature* 319, 250; 1986) until experience has shown shuttle flights to be mere routine. On the sad showing of last week, that will take some time. □



End of the road?

The British seem to be retreating from the plan to form a strong centre of clinical research.

Nearly a quarter of a century ago, after the first post-war generation of westward transatlantic travellers had returned from Europe with tales of what they had seen, the National Institutes of Health at Bethesda was almost always spoken of with that mixture of admiration and envy that presages imitation. In Britain, by a slow process, the result was eventually a decision to found a clinical research centre, under the aegis of the Medical Research Council (MRC). Now, after fifteen years, the council has been told in clear language (see p.442) that the time has come to call a halt to the venture. The Clinical Research Centre (as the institution is indeed called) has not lived up to expectations. So there are two choices: either merge it with another institution (the Royal Postgraduate Medical School, integrally part of the University of London for sure, possibly the council's own National Institute for Medical Research for another) or, the unpalatable course, close it. How can it be that a scheme that seemed a central part of Britain's research strategy a mere fifteen years ago should now be faced with the prospect of going out of business?

Nobody is particularly to blame, although it remains a fact that the centre has never functioned as what it was intended to be, a national centre of excellence in clinical research. Some of the reasons are peculiarly British. Thus, as it now emerges, those who in 1959 were envious of what was happening at Bethesda, Maryland, also took the view that "in this country the idea of a hospital purely for research is unlikely to be acceptable". Thus was born the homespun notion that the new British centre would attempt to weld an excellent research enterprise onto a general hospital, one serving the needs of a local community even if built from scratch for demographic reasons. The expectation that such a hospital could nevertheless become the focus for referrals from the rest of Britain has not been fulfilled; the London teaching hospitals are fierce competitors for patients, teaching hospitals elsewhere especially so. The result, it now appears, is that a third of the special allocation of beds and the general hospital to the centre are, in practice, in the gift of the local public health authority, not the researchers. So perhaps nobody is much to blame for the lacklustre quality of the centre's achievements over the past fifteen years. Even in its own field, other clinical centres increasingly call the shots.

In the circumstances, radical change is obviously desirable. MRC and its advisers appear to agree on that. Indeed, the report which the council has now endorsed describes the outcome of the proposed merger in glowing terms. From the ashes of past disappointment, the story goes, will rise the centre of excellence that people dreamed of decades ago. What are the chances? There are two serious impediments to the proposed recreation of the ideal of the clinical centre, of which the most immediate is boringly administrative. The past few years in Britain have shown how notoriously unwilling are institutions to give up autonomy. Medical institutions seem especially obdurate in resisting change, at least when it is imposed from outside. Forcing even two of the three institutions now canvassed into one will be an uphill task, likely to be accomplished only if some institution can bless the merger with funds on a scale no longer available. The research council is powerfully placed, knowing that all those concerned are in one way or another in its pay. But it will make no sense to create another organization to pursue excellence in clinical research on a shoestring. Better to close down the Clinical Research Centre than to force it into an impoverished marriage the outcome of which will be yet more disappointment.

The second impediment to the changes now canvassed is qualitatively different and much more serious. Those who recommend change at the Clinical Research Centre skirt around

the question why it is so difficult, not merely in Britain but in much of the rest of Europe, to persuade physicians to take research more seriously. This is not to say that there is no worthwhile activity in the field, nor centres of some distinction. But it appears to be a sad property of European medicine that the generality of physicians are more concerned with practising their craft than with turning it into a science. Even the physicians working for the health services alongside the staff of the clinical centre appear not to have been enthused, which is another proof of the disappointment of the past 15 years. Off and on, universities, sometimes with the help of charitable foundations, have sought to tempt physicians into science by one means or another, but without much success. It could be that MRC would more easily attain the objective it has been nursing all these years if it spent some of its resources on tackling the problem at source. Making clinical research worthwhile would be a start. □

Giving bugs the vote

The row about the use of genetically engineered bacteria should be stopped.

THE public inquiry (see p.441) about the proposed trial of the efficacy of *Pseudomonas* bacteria in protecting plants against frost damage has become a kind of pantomime. The idea, now well understood, spring from the knowledge that the naturally occurring bacteria are of two kinds, one of which, lacking a protein that provides a nucleating centre for water vapour, can safeguard plant cells from damage caused by ice crystals forming within them. In the wild, the proportions of the two strains of the effective species vary naturally, although there remains some doubt about the reasons why this should be so, which amounts to saying that it is not fully clear what benefits the bacteria derive from their presumably symbiotic relationship with plants. The scheme for protecting plants against frost entails the release of bacteria in which the proportions of protective strains have been artificially increased. This is the prospect described to the Monterey inquiry last month, by the ubiquitous Mr Jeremy Rifkin, as the latter-day equivalent of the first explosion of a nuclear weapons in the atmosphere.

Mr Rifkin must know that he has fertile soil to work. It is, perhaps, a pity that this first case (which, in one form or another, has dragged on for the best part of two years) did not first arise in Iowa, or Texas, where the democratic tendency to accord every possible opinion a semblance of equal time is less well developed. The difficulty in California, which the Monterey inquiry seems to have illustrated, is that freedom of expression has been extended so far that, occasionally, Mr Rifkin appears to be speaking for bacteria, not people. Cleverly he seems to have succeeded in making it possible for ordinary grown-ups to identify with micro-organisms.

The trouble, as with most of the events with which Mr Rifkin is concerned, is that the consequences of the continuation of the pantomime are potentially serious. How so? The economic case for protecting plants from frost cannot easily be denied, but is not as strong as would be, for example, a piece of genetic engineering providing all plants with a means of fixing nitrogen, even in temperate climates. (Monterey enters not because it is prone to frost, but because it is a good site for testing bacterial survival.) But the trouble with what has been happening at Monterey is that Mr Rifkin and his friends may be able to win their point by appealing to simple unreason. Although there is no basis of any kind to suspect that the end point (after a few weeks) of the bacterial populations to be released in the trials would be any different from the naturally varying populations, the doctrine of the equality of all possible opinions, which is too often mistaken for a restatement of the Bill of Rights, may carry the day. It is high time that the seriousness of the threat, not from Mr Rifkin but from the infection of the imagination that he seems bent on producing, is recognized and countered. □

US space research

Nobody knows what will happen next

Washington

THE destruction last week of the Challenger space shuttle and the subsequent suspension of the whole US shuttle flight programme has thrown into disarray what was to have been the most ambitious year yet for space science. The thirteen shuttle flights planned for the rest of 1986, including four military missions, now face an indefinite delay. Some will miss launch windows that will not recur for more than a year.

The National Aeronautics and Space Administration (NASA) is setting up an exhaustive accident investigation in order

SDI on campus

Washington

A NEW study of research contracts let by the Strategic Defense Initiative (SDI) organization highlights the Pentagon's growing influence on basic research at US universities. The Department of Defense (DoD) now supports 16 per cent of all federally funded research, up from 10 per cent in 1980. DoD represents the fastest-growing source of research support and now exceeds support from the National Science Foundation, if off-campus research is included.

The new study, by the New York-based Council on Economic Priorities, shows that more than half of federally funded work in mathematics and computer sciences is supported by DoD, as is 82 per cent of astronautical engineering and 56 per cent of electrical engineering.

The trend is likely to continue, according to the study. The SDI organization's innovative science and technology office, which targets university research, is planned to account for 5 per cent of the SDI budget; its 1986 budget is \$100 million but could expand to \$300 million by 1988. The study voices concern that, despite the SDI office's insistence that basic research will remain unclassified, secrecy clauses will be invoked as soon as the research starts to yield results of military significance.

Of SDI contracts let to universities, the great majority by value have been to Massachusetts Institute of Technology (MIT). Including the off-campus Lincoln Laboratories, MIT has so far received over \$59 million from SDI. Other major university recipients of SDI contracts are the University of Texas (\$5.6 million), Georgia Tech Research Company (4.5 million) and Johns Hopkins University (\$2.9 million).

Tim Beardsley

to discover what caused the explosion; with hundreds of photographs and abundant wreckage being discovered, as well as telemetry data intact on the ground, there seems to be a good chance that a definite cause will be found.

NASA personnel are avoiding public speculation about the cause of the accident, but it is most likely that something caused a failure of the hydrogen containment. Photographs released by NASA last weekend appear to show that the right solid-fuel booster rocket burnt through its casing about one quarter of the way up from the nozzle at least 14 seconds before the final explosion. Flames from the breach may well have played on the large external fuel tank containing liquid hydrogen and oxygen, causing it to explode.

How long the remaining three orbiters will remain grounded is anybody's guess. While some communications satellites might conceivably be switched to the European Ariane launcher (although Arianespace has full order books for the next two years), the scientific missions are designed in such a way that they have to be flown on the shuttle.

When it exploded, Challenger was carrying in its payload bay a low-cost satellite called Spartan-Halley that was to have observed ultraviolet emissions from comet Halley at its closest approach to the Sun, and a tracking and data relay satellite (TDRS) that was to be the second of what will eventually be four NASA satellites in geosynchronous network for low Earth orbital operations. The first TDRS was launched in April 1983, and a third was to have been launched next July. The TDRS system will eventually replace NASA's network of groundstations throughout the world, some of which have already been transferred to deep space uses.

The TDRS system was to have been used (though incomplete) for several planned space science missions during the year. An ultraviolet observatory known as Astro was to have been flown by the orbiter Columbia in March, with one of its principal objectives being to observe Halley at the same time as the European Giotto and Soviet Vega spacecraft intercept the comet. Dr Arthur Davidsen of Johns Hopkins University, principal investigator of one of the Astro instruments, said last week that he had not given up hope that Astro might fly in time to make observations of Halley — any time before the start of May — but admitted that the chances were slim.

Davidsen said it was particularly unfor-

tunate that the accident had occurred just before the first ever shuttle flight devoted exclusively to scientific observations, rather than to engineering demonstrations relating to science. The loss of one TDRS would not have affected Astro, but NASA officials said there is "no chance" of Astro flying in time to see Halley anyway.

The loss of one of the TDRS system satellites will have important consequences for the Hubble Space Telescope, if it is now launched before the system has two functioning satellites. The telescope had already slipped from its scheduled August launch spot to October before last week's accident, in order to give increased "contingency time" for its journey by sea from California to Florida through the Panama canal. When it will now be launched is unknown, but an incomplete TDRS system would limit data flow rates significantly, if not enough to be a "major deterrent", according to NASA.

Two major science missions that are certain to be affected are the joint NASA/European Space Agency solar polar Ulysses mission, and NASA's Galileo Jupiter probe. Both were to have been launched in May, by Challenger and Atlantis respectively, and both have a launch window that will not recur for 13 months. If neither is flown in May 1986, as now appears quite likely, then both may be competing for the same launch window 13 months later. The maximum separation of the two launches is about five weeks, which is not enough time for one orbiter to be refurbished for a second flight, and with one less orbiter in the fleet, either Ulysses or Galileo may have to wait for more than two years.

The list of space science missions that might be affected depends on how long the shuttle fleet is out of service and how NASA rearranges flight priorities. The Earth Observation Mission scheduled for launch in August must now also be in jeopardy; NASA programme managers at Marshall Space Flight Center are putting all their effort into the investigation of the Challenger explosion and have not yet studied its effects on science missions.

The Department of Defense will also have its nose put out of joint by a long grounding of the shuttle fleet; three classified payloads were to have been flown this year, in addition to a July flight including Teal Ruby, an infrared tracking experiment with applications to the Strategic Defense Initiative (SDI), and Cirrus, a US Air Force Geophysical Laboratory experiment on infrared emissions from Earth's aurora.

Further SDI experiments had been expected in 1987, and pressure from the Pentagon for shuttle payloads to make up for lost time once the fleet is back in operation will doubtless be intense; the Pentagon can "bump" civilian payloads if it chooses.

Tim Beardsley

Ariane

France waits in the wings

THE loss of the space shuttle Challenger last week has caused some heart-searching in France about the future of Hermes, the French-inspired project for a miniature space-plane to sit atop a magnified version of the European space launcher, Ariane. But according to Frédéric d'Allest, director-general of the French space agency CNES, man's presence in space is essential, so the development of Hermes will certainly continue.

The French plan has not yet been approved by the other European partners in Ariane, notably West Germany, Britain and Italy. Hermes would carry mainly people, with a small amount of cargo (4,500 kg compared with the shuttle's 20,000 kg). Heavier cargo would be carried separately, but both would be lifted into space by Ariane-5, a liquid hydrogen-fuelled version of Ariane now under development. Although safety has of course been a major consideration, critics now point out that Hermes could be more dangerous than the shuttle, because it would have only one engine.

Nevertheless, given enough warning, Hermes could use an auxiliary escape rocket to fire the space-plane upwards and out of danger more quickly than could the shuttle sitting astride its main fuel tank.

In fact, more than two years ago, Hermes designers at Aerospatiale, one of the principal French aerospace companies, were designing the escape rocket for Hermes so that it functions in what seems to be an exact analogue of last week's accident, the explosion of the second stage at the moment of maximum dynamic pressure. The designers concluded that a small rocket under the body of Hermes, capable of giving it an 8 g vertical acceleration, could save it.

This week, French space specialists would draw no firm conclusions from the shuttle disaster. On present plans, Hermes' first manned launch will take place in 1995, on what will be only the

third test flight of Ariane 5, a prospect that might now appear foolhardy. According to d'Allest, the dangers of entering space can be compared with those of entering a radiation zone in a nuclear reactor. The comparison may be more than apt, as increased public pressure for radiation protection has led to improved safety systems and correspondingly increased nuclear construction costs; similarly, the public unacceptability of an accident like last week's may well do the same for Hermes and other planned space-planes.

In the now long-running competition between the shuttle and present versions of Ariane, the commercializer of the European launcher, Arianespace, was not prepared last week to comment on the impact of the shuttle disaster. On the face of it, it would seem that satellite companies might now favour unmanned launchers, particularly as there is now a long queue of satellites waiting to be launched in what was to be the shuttle's busiest-ever year. But satellites that have been prepared for the shuttle would have to be modified to make them suitable for launching on Ariane or similar vehicles. In any case, capacity on Ariane is almost entirely taken up with a £1,000 million backlog of contracts. Although a second launch site is being opened at Kourou, French Guiana, next month, Arianespace foresees no more flight opportunities in 1986, just one in 1987 and three in 1988.

Moreover, Ariane has its own difficulties. The last flight (V15) led to destruction as the result of a leaking fuel valve. With that problem now supposed solved after a slight design modification, Ariane V16 is sitting on site in Kourou awaiting a replacement for a water cooling tank, which was found to be cracked and leaking just before the planned launch on 11 January. V16 should now fly, carrying the first French Earth observation satellite SPOT and a Swedish scientific satellite Viking, on 21 February.

Robert Walgate

Marine satellites

Japan offers free data

Tokyo

DATA from Japan's marine observation satellite MOS-1, due to be launched next year, are available for anyone willing and able to participate in the verification programme for the satellite. Japan's National Space Development Agency (NASDA) will supply MOS-1 data free of charge provided the data are used only for peaceful purposes, are valuable from a scientific and technical viewpoint and are "useful to NASDA", which means helping in evaluation of the radiometric and geometric performance of the satellite. Such is the gist of a guideline distributed by NASDA to organizations in 11 countries, including the European Space Agency (ESA). The deadline for proposals to NASDA is the end of February.

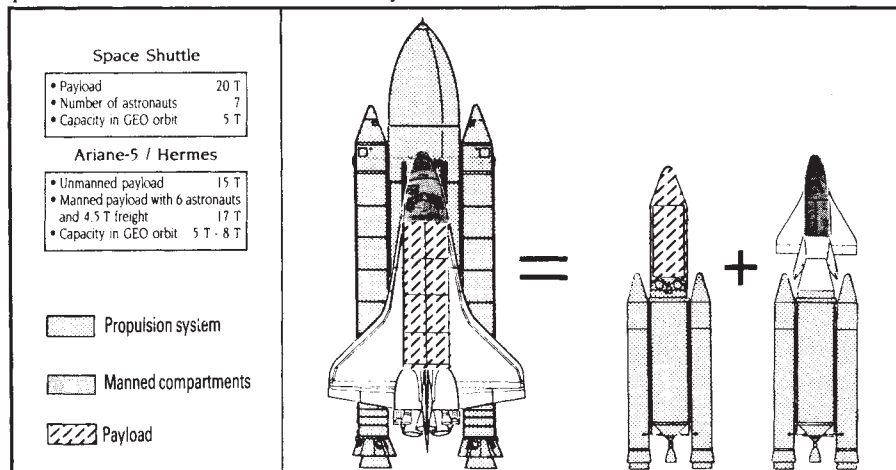
MOS-1 will view the whole surface of the Earth every 17 days and, with its three scanners and data-collection system is, theoretically, capable of transmitting more information than the Landsat satellites now in use. But coverage will be restricted by the availability of suitably equipped monitoring stations and the battery power of the satellite. Apart from NASDA's Earth Observation Center in Japan, which can monitor the satellite over north-east China, Taiwan, Korea, Japan and the eastern Soviet Union, agreement has just been reached with Thailand to build a monitoring station on the outskirts of Bangkok at a cost to Japan of 800 million yen (about £3 million). This station will cover the Philippines, Indonesia, east India and western China.

NASDA also hopes that ESA's station at Las Palmas in the Canary Islands will monitor the satellite and that Australia may upgrade its satellite stations to receive MOS data. But NASDA officials stress that MOS-1 is only experimental, with limited power. The satellite's power should be adequate if data are transmitted only during the day, but night-scanning will drain its batteries, according to Yasushi Horikawa, director of NASDA's Earth Observation Satellite Program.

MOS-1 scanners will be trained largely on the sea where its high resolution (50 m) multi-spectrum electronic self-scanning radiometer and visible and thermal infrared radiometer will provide information on ocean currents, red tides, pollution and volcanic eruptions, while its microwave scanning radiometer will monitor atmospheric water vapour content, sea ice and snow distribution.

By the time it is launched early in 1987, MOS-1 will have run up a bill of ¥56,000 million (£200 million). Its expected mission life is two years.

David Swinbanks



Voyager 2

Uranus is something else again

Washington

A NEW kind of atmospheric emission that has been called "electrogrow" was one of the most striking results of the Voyager 2 flyby of Uranus, according to Dr Ed Stone, project leader. Electrogrow, observed by Lyle Broadfoot, head of the ultraviolet spectrometry team, is believed to result from a different process than either aurora or airglow. Aurorae are caused by the precipitations of charged particles in the polar atmosphere. Airglow is the result of high-frequency solar photons in the atmosphere. Electrogrow also needs solar photons to ionize constituents of the atmosphere, but thereafter some unknown mechanism seems to provide energy to the electrons to cause the glow. The glow is observed only on the dayside of Uranus, in the ultraviolet; it had been observed before on Saturn and Jupiter, but only at Uranus did it "become clear that this was a totally new phenomenon", according to Stone. In comparison with aurorae, electrogrow is caused by very-low-energy electrons very high in the atmosphere.

The uranian magnetosphere contains mainly protons, rather than helium nuclei, for which the uranian atmosphere is the most likely source, according to S.M. Krimigis. There is in addition a warm plasma of a hundred thousand degrees and a very hot plasma of ten million degrees, reported by Fran Bagenal. Explanations for these observations have yet to be made.

The rotation period of the planet has been measured and is close to 17 hours, not too different from estimates made on the basis of the rate of precession of the rings. The magnetic field of about 0.25 gauss, with its peculiar 55 degree inclination from the rotational axis, will doubtless provide fodder for theorists for years to come.

Most of the moons are greyish in colour, suggesting that meteoritic material (together with water ice) is their main constituent. A reddish coloration would have argued for more hydrocarbons resulting from irradiation of methane.

Sizes and densities of the moons have been established; the densities of the four for which data are so far available are about 1.6, suggesting a composition of 50 per cent water ice and 50 per cent rocky material.

The uranian atmosphere was found by radiooccultation and other techniques to consist of about 12 per cent helium (plus or minus four per cent); the remainder is mostly hydrogen. This result, reported by Len Tyler, is consistent with solar composition and in accordance with expectations. A methane cloud layer at about 1.6 bars was detected; water and ammonia

were not detected directly but are presumed to exist at much greater depths.

It is clear that Uranus does have an internal heat source, but no estimate has yet been made of its size. Part of the evidence for this is that the poles are not warmer than the equator, as expected in a planet whose axis of rotation is in the plane of its orbit around the Sun. Winds on Uranus rotate in the same direction as the planet itself, also unexpected for the same reason.

New rings around the planet were discovered, but attempts to number them give rise to problems of definition, according to Jeff Cuzzi. In any event, material was found orbiting the planet both inside and outside the previously known rings. One of the surprises of the mission is that there

are few particles smaller than about a metre in diameter, in contrast with the rings of Saturn and Jupiter. One possible explanation is that if the uranian magnetic field wobbles, small particles might become magnetized and get swept out the ring; atmospheric drag might also provide an answer. Of the ten new moons discovered, at least two appear to be shepherd moons for rings.

Many of the questions raised by these preliminary results will be stated more carefully, or even partly answered, by the time the 30-day post-encounter report is completed, according to Stone. Now, the hundreds of visiting scientists who have been working at the Jet Propulsion Laboratory in Pasadena, California, for the past several weeks in preparation for the encounter have started to head for home. But the new data gathered should keep many of them busy for years.

Tim Beardsley

Soviet environment

Now pollution on television

THE new Soviet television series, "Problems, Searches and Solutions", in which experts discuss viewers' questions on key issues of the national economy, dealt on 24 January with environmental protection and natural resources. Since the programme was launched last November (see *Nature* 318, 593; 1985) the format has been changed; no attempt is made to create the illusion of off-the-cuff replies to telephone questions. Nevertheless, the pre-selected questions dealt with by the panel not only reflected a growing public interest in environmental matters, but also pinpoint a number of areas where Soviet policy needs to be strengthened.

The most striking feature of the programme is the large number of experts empanelled. Apart from Dr Yuri Izrael, chairman of the State Committee for Hydrometeorology and Environmental Control, it included a minister, a deputy minister, a vice-president of the Soviet Academy of Sciences and representatives of two other state committees and a government commission.

The discussion showed only too clearly that many of the most serious issues, such as the protection of small rivers, come under the jurisdiction of at least three of these bodies (State Committee for Forestry, Ministry of Land Reclamation and the Commission of the Presidium of the USSR Council of Ministers for the Protection of the Environment and natural resources), to say nothing of bodies involved in the long-term Food Programme.

Similarly, the drainage of forest marshland in Lithuania and Byelorussia turns out to involve not only the same three governmental bodies, but also a Byelorussian research institute working on the cul-

ture of heavy-fruited cranberries in natural marshland or industrial plantations. Environment could well benefit from the kind of superministry recently established for agriculture, as part of Mr Mikhail Gorbachev's drive to streamline the economy.

The Gorbachev reforms have undoubtedly made it easier to discuss past mistakes in environmental planning. Both *Pravda* and *Izvestiya* have given extensive coverage, in the past few weeks, to the environmental threat from the two cellulose combines on Lake Baikal which, in spite of expensive purification installations, still continue to endanger the lake.

Dr Aleksandr Yanshin, vice president of the Soviet Academy of Sciences, pointed out in the television programme that the pollutants discharged by these combines are incorporated in the shells of the minute crustacean, *Epischura baikalensis* which, when it dies, carries the shell to the bed of the lake. But the life-cycle of *E. Baikalensis* is short, and the bed conditions of the lake play an important part in maintaining the exceptional purity of the lake water under natural conditions; even a relatively small discharge of pollutant is a considerable threat.

The chairman of the Committee for Forestry admitted that, although over-felling had been reduced by some 80 per cent over the past fifteen years, some 12 million m³ of wood is still being over-felled every year, thus violating "the basic principle of timber use". But only a few years ago, over-fulfilment of targets was urged in every sector of the Soviet economy, including forestry. Some officials may not yet have made the necessary mental adjustment to the new circumstances.

Vera Rich

Animal welfare

OTA squares circle of dissent

Washington

THE Office of Technology Assessment (OTA) has weighed in to the debate over the use of animals in research with a report that even critics regard as comprehensive, and which will have to be reckoned with in any future debates on the subject.

More than two years in the making, at a cost of \$425,000, the report (*Alternatives to Animal Use in Research, Testing and Education*) looks into alternatives to the use of animals, not just in research, but in education and testing as well, judging

them by the extent to which they replace, reduce and refine present techniques.

While acknowledging that "some biological research requires — and in the foreseeable future will continue to require — the use of live animals", the report notes that the alternatives to tests such as the LD₅₀ test that are now more widely adopted in industry, and looks for potential parallels elsewhere in testing and research. To reduce the use of animals in science, the role of computers can be expanded beyond that of a tool for modelling animal behaviour. Using computers to improve information exchange about data already collected can eliminate redundancies in research. But the report acknowledges that the United States lags behind European countries and Japan in the protection of laboratory animals.

Difficulty in estimating the numbers of animals bedevils attempts to evaluate current usage. The best data come from the Animal and Plant Health Inspection Service (APHIS) of the US Department of Agriculture, which collects records on the use of dogs, cats, hamsters, rabbits, guinea pigs and non-human primates. This group of animals represents only about 10 per cent of animal use; reporting on rats, mice and birds is not required. Extrapolating from these numbers, OTA estimates that between 17 and 22 million animals were used during 1983 in research and testing.

Even before the report was officially released, some groups were rejecting it. The People for Ethical Treatment of Animals calls the report biased, saying that it does not reflect the views of the animal rights movement. Connie Kagan of the Animal Political Action Committee calls it "disappointing", and doubts if it will have much impact on Congress.

But most observers agree that the report is as comprehensive as could be reasonably expected, and Barbara Orlans of the Scientists' Center for Animal Welfare, believes it will be "recognized as a bible".

Sensitivity about the use of animals in science has been growing recently in the United States. The Animal Welfare Act was amended last year, strengthening its provisions, and the Health Research Extension Act of 1985 puts the force of law behind some parts of the current Public Health Service guidelines on animal use. These changes in federal law have shifted much of the responsibility for oversight of animal use onto institutional review committees, and OTA estimates that "the overwhelming majority of experimental-animal users in the United States" now come under the oversight of a "structured, local review committee".

Perhaps the single greatest accomplishment of the report is that it seems to have generated surprisingly little acrimony. According to Arthur Caplan, chairman of the OTA advisory panel, the report transcends "the name calling, finger pointing and hair pulling that normally accompanies the discussion of animal use in science".

Joseph Palca

Preaching to converts

Washington

GROWING awareness may have inspired the report of the Office of Technology Assessment (OTA) on alternatives to animal use, but in the commercial field, business savvy may provide its own reasons for abandoning animal resources. So Charles River Biotechnical Services (CRBS) joined the industry-wide shift from *in vivo* to *in vitro* monoclonal antibody production with last month's purchase of the Opticell branch of KC Biologicals. Executive vice-president Melvin Balk says the acquisition will expand the Massachusetts company's bioprocessing services to include vaccines and recombinant proteins, while Opticell's culture system will gradually usurp the mouse as CRBS's antibody factory.

Balk also served on the advisory panel for OTA's project, as much, he says, to scout for alternatives that make "economic sense" as to further OTA's evaluation. In the case of cell culture methods for monoclonal antibody production, economic sense and ethical sensitivities coincide. Balk says the techniques avoid costly problems associated with, for example, variability and contamination of mouse monoclonals. But the success of these methods is still contingent on improvements in product yield and the stability of large-scale markets. Although *in vitro* products such as human chorionic gonadotropin are aimed at high-demand diagnostic or therapeutic markets, most *in vitro* methods cannot yet equal the milligram per millilitre harvest that mice supply.

Echoing OTA's conclusion, Balk insists that mice will always be necessary for monoclonal production because some hybridomas simply will not thrive elsewhere. Other industry experts paint a more optimistic picture for the future of cell culture. Allan Jarvis at Damon Biotech anticipates a time when *in vivo* production may be superfluous and outdated; Damon has survived five years without it. At Becton-Dickson in California, the mood is more cautious. Before launching an *in vitro* programme, the company's attempts to evaluate different culture systems have been frustrated because these systems are still in their largely untried infancy. The venture will probably be risky and, as one manager remarks, a half-million-dollar mistake is not a welcome alternative to animal use.

Karen Wright

Guidelines off

Washington

THE new guidelines adopted by the National Science Teachers Association (NSTA) for animal use in US elementary and secondary schools will not be endorsed by the National Academy of Sciences.

While the Office of Technology Assessment (OTA) report indicates that the "overwhelming majority" of animals used in research and testing are governed by some sort of institutional review process, the same cannot be said for the use of animals in education, especially at the elementary and secondary school levels.

What guidance there is comes primarily from guidelines established by NSTA and the National Association of Biology Teachers (NABT).

In 1981, NSTA submitted its guidelines to the National Academy of Sciences, and they were eventually endorsed by the National Research Council. The guidelines prohibit "experimental procedures on mammals, birds, reptiles, amphibians or fish that cause the animal pain". The 1981 guidelines also state that "students shall not perform surgery".

But not all members of NSTA and NABT were content with these restrictions. Many felt that they would hamper the education of outstanding students. Moreover, the guidelines were more restrictive than those accepted by the International Science and Engineering Fair, which sets standards for many school science fairs.

So in 1985, NSTA rewrote the guidelines. The language was altered to read that no experiment causing "unnecessary pain" should be performed, and surgery was permitted, so long as it was supervised.

In December, NSTA resubmitted their guidelines to the academy, hoping for endorsement. But the academy declined, although it has yet to issue a public statement on the matter. One possibility is that the academy will issue its own guidelines later this year, and these are likely to conform more closely to the 1981 NSTA guidelines.

Joseph Palca

Pacific volcano

Transient Iwojima calf

Tokyo

THE Japanese islands burst into life last month when a submarine volcano south of Iwojima broke surface, sending a plume of smoke and ash billowing skywards. Within a day, a crescent-shaped island almost 1 km in length had formed around the volcanic vent, but, as with previous attempts to join the Japanese archipelago, the island will probably disappear again.

The submarine volcano Fukutokuokanoba lies 1,300 km south of Tokyo at a point 5 km north-east of Minami (south) Iwojima (24° 17' N, 141° 29' E). Its first recorded eruption began on 14 November 1904. By 1 February 1905, an island 5 km in circumference and 150 m high had formed and was named Shin (new) Minami Iwojima. But by June of that year, it was reduced to a shoal 150 m in circumference and, by 1910, the point where the island had been was 424 m deep.

Infrared satellite images indicate that the present eruption began on the night of 18 January. Activity peaked on 20 January when a column of smoke 50 m wide rose to 3,000 m, blocks of volcanic rock up to 10 m in diameter were hurled hundreds of metres into the air and a plume of discoloured water spread for about 20 km in a

south-easterly direction. The island was reported to be 700 m long from east to west, 500 m wide and 15 m high, but a week later it was 500 m in length and showed signs of rapid erosion.

The little island's big brother, Minami Iwojima, is a 3.5 km² uninhabited jungle-capped cone of basalt. It lies at the south-



ern end of a 1,250-km-long strip of active volcanoes that starts at Mount Fuji and then runs along the edge of the Philippine plate through the Izu Seven Islands (Shichito).

While most of the islands of the Shichito-Iwojima chain are composed of tholeiitic basalt typical of the front of volcanic island arcs, the rocks of Minami Iwojima and Iwojima (also known as Naka (middle) Iwojima) belong to the alkali suite usually found at the backside of island arcs. Rock collected from the temporary islands formed in 1904 and 1914 and from the top of Fukutokuokanoba in 1982 were all olivine-augite trachyandesite with high alkali contents. The new volcanic island is thus probably closer in composition to Naka Iwojima 50 km to the north-west than to nearby Minami Iwojima which is composed of alkali basalts, according to Shigeo Aramaki of Tokyo University's Earthquake Research Institute.

In contrast, Kita (north) Iwojima, about 70 km north of Naka Iwojima, is a tholeiitic basalt volcano. The abrupt change to alkali volcanics at the southern end of the Shichito-Iwojima ridge is puzzling. Alkali-rich magmas are thought to form at high pressures at the lower end of the plate subducting beneath the volcanic island arc. Thus, according to Teruaki Ishii of Tokyo University's Ocean Research Institute, the occurrence of alkali volcanics in the southern half of the group of islands may be related to the steep dip of the subducting Pacific plate as the Ogasawara Plateau, a large guyot riding on the Pacific plate, plunges into the Izu-Bonin trench immediately opposite the islands. Ishii points out, however, that this explanation fails to account for the lack of tholeiitic basalt volcanoes between Iwojima and the trench. **David Swinbanks**

AIDS

Academy looks for strategy

Washington

A NATIONAL agenda on research and health care for acquired immune deficiency syndrome (AIDS) is the objective of a six-month project of the National Academy of Sciences and the Institute of Medicine (IOM) in the United States; the approach so far is considered to have been random and uncoordinated.

A single report will be generated from the work of two panels, one on research and development, the other on health care and public health. According to Sam Thier, president of IOM, the report is not to be a definitive treatise on AIDS, but rather part of "a continuum of effecting national action".

The panel on research and development will examine current research efforts to determine if they are properly organized and sufficiently well supported. IOM hopes to weigh up whether a programme on the scale of the Manhattan Project or Project Apollo is justified by the size of the current health problem, and whether such a large-scale programme would provide a return commensurate with its costs.

"It is critical to be sure that for an area as important as this that the research agenda and strategy is as well crafted as we can have it", says Thier. "We have to be sure that all segments of the research community are integrated into the effort, and if necessary recruited into it where they are not as actively involved."

The other panel will investigate the social and public policy questions raised by AIDS. It is still unclear how the health-care delivery system can most effectively cope with a disease for which there is as yet no cure. Some rational strategy that takes into account the needs of the AIDS patient and the needs of society must be developed.

Almost any active steps to control the AIDS epidemic will of course cost money, a conclusion that troubles law makers in these days of the Gramm-Rudman Act and it will force budget reductions. While Thier acknowledges these financial problems, he considers the IOM effort should not have to be sidetracked by the uncertainties of funding.

"It's all well and good to say we have Gramm-Rudman, and there are cutbacks and so on", says Thier. "But that is not going to stop the progress of this disease."

The IOM report is being put on a fast track. Thier is anxious to see this and other IOM projects completed faster than the 18 months typical for reports in the past. IOM could then play a larger role in affecting national policy on the critical health issues.

Joseph Palca

Trials on ice

Washington

AFTER a lively public hearing last week at Salinas, California, Advanced Genetic Sciences (AGS) announced that it will postpone for thirty days the environmental release of its genetically altered bacteria designed to protect plants from frost damage (see *Nature* 319, 254; 1986).

Monterey County Supervisors scheduled the hearing to air questions about the safety of the release. Sam Karas, chairman of the board of supervisors, said some of the most influential evidence came from Dr Lee Cavalieri, retired member of the Sloan Kettering Institute for Cancer Research. Karas left the meeting with the understanding that Cavalieri was suggesting there is evidence that the experiment may pose some cancer risk, but Cavalieri emphatically denies making any such statement.

AGS has already won approval from the Environmental Protection Agency and the California Department of Food and Agriculture to proceed with the test on a farm near Salinas, but local opposition brought its plans up short. Despite vocal opposition, AGS plans to proceed with the trial, but not before it has satisfied the local authorities that the bacteria poses no threat to residents. A compromise may be to move the site of the trial to a less populous part of the region.

Joseph Palca

British clinical research

Shotgun merger proposed

A SINGLE centre of excellence is the only solution to Britain's national "crisis in clinical research", according to a committee reporting last week to the Medical Research Council (MRC). Although set up to advise MRC on the future of its Clinical Research Centre (CRC), the most costly of its research establishments, the committee strongly urges a merger of CRC with the Royal Postgraduate Medical School (RPMS), an entity under the wing of the University of London. Ideally, it says, the National Institute for Medical Research (NIMR) should also be merged with CRC and RPMS. And if it is not possible to combine CRC with RPMS, CRC should be disbanded.

Indications of the poor state of UK clinical research include the particularly marked fall in grant applications to MRC from clinical scientists, according to the committee. This reflects a situation where young doctors are increasingly unprepared to obtain delayed professional accreditation by undertaking research training. And established clinical

academic staff are too beset by administrative problems brought about by financial constraints to apply for grants and hold together a research team.

Chaired by Sir Michael Stoker, and taking over where his previous committee left off (see *Nature* 314, 487; 1985), the new committee reports that "CRC has not developed the national role envisaged when it was first established" fifteen years ago. While offering faint praise for CRC's director, Sir Christopher Booth, for bringing together basic medical and clinical science, the committee's report says that yet more basic science is needed, whence its suggestion that NIMR should be party to the merger.

Stoker's committee identifies as one of CRC's chief problems its attachment to a hospital that has developed no specializations on which the research could be focused. Blaming this on a failure to develop an appropriate administration of the hospital, with a greater role for the director of CRC, at Northwick Park in North London, the committee concludes that "there is a growing divergence of the two populations of clinical staff on the site"; it sees little prospect of the necessary unity of purpose emerging, and that CRC suffers from the lack of a graduate programme.

Sir Christopher Booth acknowledges both failings, but says that the report overstates the lack of unity between the hospital and the centre, and that it is misguided in its implied criticism of a programme of research based on "common" diseases — those that predominate at a general hospital — rather than the more rare disorders by which researchers are attracted.

Reactions to the suggested merger are generally favourable, at least in principle. MRC has endorsed the report, Booth strongly supports the idea of a single centre of clinical research and Sir Frederick Dainton, chairman of the council of RPMS, has said that the suggested merger is an exciting prospect.

For RPMS, the most tangible gain would be greater financial security. Cutbacks in university funds and a fall in the number of overseas students following the British government's withdrawal of subsidy for them mean that RPMS now exists on more "soft" than "hard" money according to its dean, Dr D.N.S. Kerr. CRC, by contrast, has an enviable number of established MRC posts. A merger of the best of the two institutes would make sense in those terms, particularly since MRC spends £6.3 million a year at RPMS and just over £10 million at CRC.

Doubts at RPMS focus on two points. The first is the extent to which the school's ambition for more and better basic research would be met by a merger with

CRC, which itself is lacking in that respect. Second, and perhaps inevitably, there would be no enthusiasm for moving to CRC, which has more space but not as good a relationship with Northwick Park Hospital as RPMS has built up with Hammersmith Hospital in West London. But in principle CRC has 160 beds at Northwick Park at its disposal, while RPMS has none reserved for its use.

Sir James Gowans, secretary of MRC, believes that "the barometer is set fair" for a merger of CRC and RPMS, that the merger has to be on one site and that the exercise would cost rather than save money in the short term. If no merger could be agreed, the council would "return to the recommendation" that CRC be disbanded, a view MRC has not endorsed.

As to the possibility of NIMR joining CRC and RPMS, the prospects seem remote. The original intention of MRC had been to merge CRC and NIMR, but the idea was set aside in 1980 and NIMR has since set off on a different course under a new director. Moreover, NIMR does not have access to a hospital and so would have to move to one of the other two sites at enormous cost.

Peter Newmark

MRC attacked

THE British Medical Research Council (MRC) is under attack in the pages of the *British Medical Journal* for its lack of support of overtly medical research.

The correspondence was started by Professor Miles Irving of Hope Hospital in Salford who criticized the council's intention to close down its trauma unit in Manchester. He implies that MRC's failure to find a suitable replacement for the retiring director reflects the council's lack of understanding of the importance of such research. "To many surgeons and anaesthetists the council appears to be out of touch with everyday clinical practice".

Support for this view comes from Professor John Dobbing of the University of Manchester, who claims that his own research on the effects of social and nutritional adversity on the development of brain and behaviour was denied further support by MRC at the same time as being praised in its annual report.

In replying to Irving, Sir James Gowans, secretary of MRC, denies any bias against surgical disciplines, a view in which he is supported by a member of one of MRC's boards, Professor David Hamblen of the Department of Orthopaedic Surgery at the Western Infirmary in Glasgow. "Until surgeons regard research as providing a scientific basis for all surgery and not just occupational therapy for academic departments", he writes, "the situation is unlikely to improve."

Peter Newmark

Budget of INSERM grows

THE French medical research council INSERM, whose generous budget has always been the envy of biologists supported by the larger but multidisciplinary research council, the Centre National de la Recherche Scientifique (CNRS), will have four per cent more in real terms to distribute in 1986 than in 1985.

The new money will be spent, INSERM says, on fifteen new laboratories, on improved support of existing laboratories, a doubling of the budget for large laboratory equipment and on 70 new research posts, a four per cent increase on the present INSERM research staff of 1,700. There will also be 20 new temporary posts for foreign researchers.

The new research groups mostly fall in the "priority research areas" identified by INSERM. Grants to extramural groups in university, CNRS, hospital and other laboratories will also increase this year, by a substantial 19 per cent. But the FF50 million (£10 million) spent outside INSERM's walls is less than a tenth of the FF519 million (excluding salaries) that will be spent in the council's own laboratories in 1986. Including salaries and overheads, INSERM's gross budget for the year will be some FF1,650 million for 1,700 scientists. By comparison, the budget of the CNRS life sciences department is similar, at around FF1,800 million, but it must support some 2,700 researchers, whence the envy INSERM excites. Robert Walgate

Bulgar identity

Proofs of ethnic purity advanced

A MAJOR research project into the present state and future prospects of Bulgarian brain and intellect, carried out at the Institute of Brain Research of the Bulgarian Academy of Sciences, suggests that the modern Bulgarian has a very flexible intellect, but fails to use it to its full potential. According to Dr Ana Vrbanova, the director of the institute, this unsatisfactory situation is caused by poor use and organization of Bulgarian culture and traditions.

EPA expansion urged

Washington

AN advisory panel to the US Environmental Protection Agency (EPA) has concluded that EPA needs to expand its research programme in order to cope with the expected influx of requests to release genetically altered organisms into the environment. The panel says it is impressed with what EPA has done, given the resources at its disposal, but says the agency should do more to establish test protocols for considering environmental releases.

The Study Group on Biotechnology, set up under EPA's congressionally established Science Advisory Board, is a public group intended to provide external review of the agency's actions. The study group was chaired by Martin Alexander, a microbiologist at Cornell University. EPA has not officially responded to the report, but its scientists say they broadly concur with its conclusions. EPA recently approved the first environmental release of genetically engineered microorganisms into the environment, and ran straight into a controversy (see *Nature* 319, 254; 1986). The agency will shortly be publishing general terms and conditions that it will use for considering such applications.

The study group says an adequate research programme for environmental releases should address survival, growth, genetic transfer *in situ*, dispersal, environmental effects, health effects, remedial action and use of biotechnology for destroying pollutants. So far, EPA has examined applications on a case-by-case basis but the study panel says the "absence of generally accepted tests and test methods . . . is cause for concern". The panel observes that EPA has "essentially no program" of research into environmental effects, and says EPA needs increased resources to conduct "innovative research" that will lead to definitive generally accepted test protocols. It also recommends that another special panel be established to provide peer review of EPA's risk-assessment efforts. Tim Beardsley

Reporting on the preliminary results of what is admitted to be a far from complete study, the Bulgarian media have stressed the declaration of a recent Central Committee plenary meeting that "the construction of a developed socialist society not only asks for but also demands the activation of the brain cells".

This premature media coverage is not simply a way of urging Bulgarians in the learned professions to work harder during the new five-year plan, but seems to be related to current political and ethnographic claims of the purity of the Bulgarian nation and, in particular, of the absence of a Turkish ethnic minority.

This emphasis on Bulgarian specificity accords well with the current official Bulgarian view of its Turkish minority. During the past few years, large numbers of Bulgarians of Islamic descent have ("voluntarily", it is claimed) changed their Turkish-sounding given names and surnames to Slavonic names. Protests from Turkey that the changes were being carried out under pressure brought a rejoinder from Sofia that the people concerned were not Turkish at all, but ethnic Bulgarians whose ancestors had been converted to Islam under the Ottoman Empire.

Now the Bulgarian Academy of Sciences has published a monograph, *Anthropology of the Bulgarian Nation*, said to be based on more than 30 years work by the Institute of Morphology of the Bulgarian Academy of Sciences, using, in particular, anthropometric measurements carried out in areas hitherto thought to contain members of Bulgaria's Turkish and Greek "minorities".

According to a review in the daily *Otechestven Front* (1 January 1986), the studies revealed so high a degree of homogeneity that the experts conclude that the fusion of Slavonic, Thracian and Asiatic ethnic elements that gave rise to the Bulgarian nation was complete by the end of the tenth century, and that the past 1,000 years of foreign incursions, including centuries of Ottoman rule, have left no ethnic trace whatsoever. Bulgaria's Turkish minority, the reader is left to deduce, does not exist.

This remarkable piece of dialectic is not Bulgaria's first effort in political anthropology. In 1978, the Bulgarian Academy produced a monograph *Macedonia*, which set out to prove that the Macedonian nation did not exist.

The international repercussions of the academy's new monograph too are likely to be acrimonious, particularly in Greece and Turkey. The study of Bulgarian brain-power, although clearly part of the same political stance, will probably have its greatest effect at home, and may possibly

engender conflict within the Bulgarian scientific establishment. For, according to Dr Vrbanova, the main purpose in studying Bulgarian brain-power is to activate sociologists, psychologists and doctors into finding and developing techniques to expand the potential of "Bulgarian's nine million brains".

Vera Rich

Marine science

Changes afoot

PLANS are afoot for changes in the style of British marine research. Even as the House of Lords Select Committee on Science and Technology finished its report (see *Nature* 319, 351; 1986), the Science and Engineering Research Council (SERC) was preparing to divest itself of its Marine Technology Directorate (MTD). With a 1984-85 budget of more than £5 million, it was established over nine years ago to develop academic involvement in marine matters. It achieved this by a system of grants to seven centres mainly in British universities. Within the next few months, the directorate is to be replaced by a separate corporate entity, MTD Ltd, under the aegis of the Fellowship of Engineering (the engineers' equivalent of the Royal Society). SERC will for the next three years provide only one-third of the £300,000 operating costs, with the balance being raised by subscription by industry and government departments. SERC will also cut its research contribution and hope that industry will plug the gap.

SERC policy is that directorates should be discontinued as in-house activities after a formative period. A.M. Adye, director of MTD, said that he was not aware of any change in SERC plans in response to the House of Lords committee recommendation that an overseeing Marine Board be established within SERC. But he agreed with the main thrust of the report — that there is too much fragmentation of interests and an overview is needed.

Hugh Fish, chairman of the Natural Environment Research Council, supporting his council against the attacks contained in the report, also agreed that greater co-ordination was necessary. But he felt that a better solution could be found than "the bolting-on of a Marine Board . . . to the existing organizations".

Another spoke in the wheel is the Department of Trade and Industry's Marine Technology Committee, which handles the department's marine research budget. This body is just starting the ball rolling in an effort to develop a co-ordinated national programme for ocean resources such as metallic nodules and "soft" energy (excluding hydrocarbons). The main aim is to encourage industrial interest in both these resources and the technological infrastructures needed for their exploitation.

Peter Gambles

Remote viewing exposed

SIR—Some years ago, we challenged experiments on remote viewing conducted by Targ and Puthoff¹ on the grounds that sensory cues were available to the judges, enabling correct matching of the subjects' descriptions against the target locations². This objection was countered when another investigator, Tart, attempted to remove all the cues and organized a rejudging that yielded results as significant as those originally obtained³. At this stage, we asked for access to the data, and we are now pleased to acknowledge that, after a three-year interval, Puthoff has released the relevant data from the first experimental series with the subject Price. Having examined both the edited and unedited transcripts, we note that Tart failed to remove a number of potentially useful cues from the transcripts given to his judge. This bias in Tart's editing, as was suspected earlier⁴, invalidates the judging exercise.

The Price series of remote viewings consisted of nine experiments. Eight of the allegedly edited transcripts (all except the one from Experiment 2) retain information concerning the subject's location. These were a park in Experiments 1 and 2, an office in Experiment 5, and a shielded room in the remainder of the experiments. The transcript for Experiment 1 also retains the passage at the very beginning where Price expresses apprehension about being able to perform the task: "the feeling that one can't or won't be able to do it and shouldn't even try it". The transcript from Experiment 3, the first to be conducted inside the shielded room, quotes a question from Targ as to whether the subject noticed "any difference in being in a shielded room compared to being in the park".

Targ and Puthoff's publication *Mind-Reach*⁵ contains extracts from the transcripts for Experiments 1, 4 and 9 and publishes in full the transcript for Experiment 7. *Mind-Reach* also lists the correct order of target sites and the subject's location during each of the experiments. Given the public availability of *Mind-Reach* at the time when the rejudging was conducted, it was not a valid procedure to include all nine of the transcripts, as the judge could easily look up the relevant information or recall it using memory based upon earlier reading of the material. Only the five targets not cued by the information in *Mind-Reach* should have been employed (Experiments 2, 3, 5, 6 and 8) as in the case of the unsuccessful rejudging reported by Marks and Kammann² which used transcripts with all cues deleted. Furthermore, the cues contained in the transcripts associated with Experiments 1 and 3 invalidate the inclusion of Experiments 1, 2 and 3 as the transcripts

for all three of these experiments can be readily matched using the available cues. Experiment 5 is also invalidated because its transcript contains the uniquely diagnostic cue of the subject's location (an office) used only on this occasion. Hence only Experiments 6 and 8 remain potentially uncued in the rejudging exercise.

Considering the importance for the remote viewing hypothesis of adequate cue removal, Tart's failure to perform this basic task seems beyond comprehension. As previously concluded², remote viewing has not been demonstrated in the experiments conducted by Puthoff and Targ, only the repeated failure of the investigators to remove sensory cues.

DAVID MARKS

Department of Psychology,
University of Otago,
Dunedin, New Zealand

CHRISTOPHER SCOTT

60 Highgate Hill,
London N19, UK

1. Targ, R. & Puthoff, H. *Nature* **251**, 602–607 (1974).
2. Marks, D. & Kammann, R. *Nature* **274**, 680–681 (1978).
3. Tart, C.T., Puthoff, H. & Targ, R. *Nature* **284**, 191 (1980).
4. Marks, D. *Nature* **292**, 177 (1981).
5. Targ, R. & Puthoff, H. *Mind-Reach* (Delacorte, New York, 1977).

No sexism please

SIR—Gerald Weissmann, in his zeal to defend the University of Texas Health Services Center (*Nature* **318**, 308; 1985), appears to have committed a gaffe of his own. Presumably "daughters" as well as "sons" of the university will be future Nobel laureates, assuming that the university employs both male and female scientists.

ANN S. HOBBS

Department of Physiology,
University of Maryland,
School of Medicine,
660 West Redwood Street,
Baltimore, Maryland 21201, USA

Future UNESCO

SIR—Whatever the reasons for the withdrawal of the United States and Britain from UNESCO, such drastic action is needed in my opinion to dismantle an institution that has sunk into a vicious circle of rot. UNESCO is not the only UN agency that needs stripping; others also suffer innumerable faults in their operation. The United Nations embodies the highest ideals of mankind and if its scientific agencies are to function with minimum errors and deviations, then they should be serviced by the highest calibre of scientists, not by those whose prime motives are in financial reward and political glory.

Your leading article (**318**, 397; 1985) omits a dire consequence that may result from the departure of Britain from

UNESCO — the export of inferior values in science and education to the English-speaking developing countries in Africa by those who are probably eager to step into the niche left by the departing British. Modest but good and useful science was established by the British in their former colonies. The high quality of relevant environmental sciences in East Africa before and for about a decade after the independence of these countries is one such example. Subsequently, scientific research in these countries declined. Now, I cannot help thinking of the possible infiltration into these countries of what I see as "science without its scientific ethics", and the threat that this would pose to the further decay and distortion of the image of science in these countries.

W. OKERA

43 Margaret Street,
Walkerville,
South Australia 5081

Catastrophism

SIR—I refer to Paul R. Weissman's letter (*Nature* **316**, 572; 1985) "Catastrophism still unexplained", and the 1978 and 1979 references.

The suggestion that the Sun's motion in the Galaxy might be a possible cause of terrestrial extinctions was made by J. Steiner as early as 1967¹ and again, independently, by A.A. Meyerhoff in 1973², though these references suggest the Sun's intragalactic orbit as the cause of the 250-million-year cycle if it exists.

W. STANNAGE

5 Coastguard Station,
Robin Hood Bay,
North Yorkshire YO22 4SY, UK

1. Steiner, J. *J. geol. Soc. Aust.* **14** (1), 99 (1967).
2. Meyerhoff, A.A. *Can. Soc. Petrol. Geol. Memoir* **2**, 745 (1973).

Eureka!

SIR—Buckminsterfullerene is not only a clumsy name for an outstandingly beautiful molecule (*Nature* **318**, 162–164; 1985); it is also historically inappropriate. The truncated icosahedron was known to the First Century Alexandrian mathematicians Heron and Pappos, who refer to a text on this and a dozen other such solids by Archimedes. The original treatise has since been lost, but it would be a fitting tribute to call the new substance Archimedene. This could be abbreviated to Archene (particularly suitable if this should indeed turn out to be one of the oldest molecules in the Universe — Greek *arkhe* 'origin'), or ARCH (all-round carbohedron).

PHILIP J. STEWART

University of Oxford,
Pauling Human Sciences Centre,
58 Banbury Road,
Oxford OX2 6QS, UK

Galactic voids may be statistical

Excitement in the past few years about the apparent clumpiness of galactic clusters, leaving empty spaces in between, may be misplaced.

WHEN is a void real and not just a statistical artefact? The question has become topical and important because of recent reports that the Universe, statistically homogeneous though it may be, is also clumpy. There is no preferred direction in space, at least within the limits of experimental errors as they are, but nevertheless the visible matter is not merely collected into galaxies but the galaxies themselves appear to occur as clusters. More than that, as well as clusters there appear also to be voids, patches on the sky where there are no recognizable clusters of galaxies. Is this a phenomenon requiring a physical explanation? Or may the appearance of the voids be a statistical illusion which does not contradict the notion that galaxies and clusters thereof are distributed at random?

The circumstances are, or should be, familiar to cosmologists and astrophysicists. Observational bias is a recurring trouble. Whenever new techniques are used to search the sky, there is obviously a risk that the first novel objects to be found will have declared their presence only because they are exceptional. The only remedy is patience, but that is not as simple as it sounds. In spite of all the energy expended on the observation and interpretation of quasars in the past quarter of a century, for example, it is still not certain whether the apparent scarcity of quasars at a distance greater than that corresponding to a redshift parameter $z=3.5$ or thereabout implies either that quasars do not often form in more distant or younger galaxies or that there is an observational bias against the detection of more distant objects. The quasars are also the cause of a running battle between the US observer Arp and most other quasar watchers about his assertion that these objects turn out to be more often geometrically related to each other, in straight lines for example, than chance would allow.

The treatment of the problem of voids by David H. Politzer and John P. Preskill of the California Institute of Technology (*Phys. Rev. Lett.* **56**, 99; 1986) should be read by all those embroiled in arguments like these. Briefly, the conclusion is that the voids so far described on the surface of the sky are almost as likely to be consistent with randomness as with some physical process for concentrating galaxy clusters elsewhere, away from the voids. The curious feature of this article is that the underlying problem is so important and

general, and the solution so simple, that one is bound to wonder why it has not been dealt with in a quite different context. At the very least, it will have cropped up as a problem for students to solve in a final examination at one of the better Victorian universities.

At first sight, indeed, there seems to be no problem. For the sake of definiteness, the case of the two-dimensional observable surface of the sky is a good place to start. The objects observed may be galaxies, or galaxies brighter than some arbitrary limit, or clusters, or clusters with redshifts between arbitrary limits. "Voids" will be only relatively empty; a patch of sky of area S will not be entirely devoid of objects, but, rather have fewer, perhaps many fewer, than most other patches of the same size. So the problem seems tractable by the elementary application of Poisson statistics. The chance that a patch of area S will contain k objects will be given by $(1/k!) (nS)^k e^{-nS}$, where n is the average density of objects, just as it is the statistical distribution of the numbers of soldiers in the old Prussian army killed each year by being kicked by a horse. A void corresponds to the cases in which k is less than the average number in a patch of area S , given by nS .

The snag, unfortunately, is that a calculation that a specified patch in the sky will be a void, however defined, does not directly correspond to what people with telescopes actually observe. Physically, what matters is not that a specified patch on the sky should contain a void, but that the sky as a whole should contain some such patch. So the problem that matters is to calculate the probability that a void with predetermined shape and content will appear somewhere in the sky from the assumption that the objects being observed are distributed randomly. The trick, perhaps not surprisingly, has something in common with the process of swinging a telescope across the sky, searching for a void with characteristics specified in advance.

The principle is simple even if the algebra may be tedious. If a void is defined in advance as a circular patch that must contain no more than, say, k objects, the sensible course to follow is to cover the sky with a regular array of overlapping patches of that size, calculating the chance that one or more of them may correspond to a void, and then to allow the density of the overlapping patches to increase inde-

finitely. Technically, this is done by centring circles on the intersections of some kind of lattice structure whose spacings are then made indefinitely to shrink. The same kind of calculation can be carried through with other shapes of voids than circles.

Not surprisingly, the chance of finding a specified void somewhere in the sky turns out to be greater than that a particular patch will contain such a void. More significantly, numerical factors apart, the chances of finding predetermined voids are related to the Poisson distribution by factors of the form (nS) raised to the power two in the case of circular patches on a two-dimensional sky, but which may be raised to a higher power if the voids are elliptical in shape and allowed to assume any orientation. What this implies is that the presence of voids in the distribution of objects in the sky is certain to be much more common than has been assumed.

The argument applies directly to one of the first claims of the presence of a void in the distribution of rich galaxy clusters (Bachall, N. and Soneira, R. *Astrophys. J.* **262**, 419; 1982), which has been the stimulus for other such claims in the more recent literature. As Politzer and Preskill argue, the chance that the observed distribution has arisen from a random distribution of galaxies is something like 30 per cent, which goes to show that the occurrence, dramatic though it may appear, is nowhere out of the ordinary.

There is, of course, no reason why anybody should be upset by this turn of events. What stands out a mile is that simple expectations of the consequences of the random distribution of objects on the sky are unreliable. The surprise is that the accurate calculation has not been made before this. And even if it should be that voids are less remarkable than they have seemed in the past few years, that does not imply that they are entirely without interest.

For galaxies are not distributed at random. Large clusters are gravitationally bound and presumably collapsing. Much of the immediate interest in this field is that of measuring the dynamics of the clusters, which may throw light on their origin. This why Politzer and Preskill want to turn their problem around, and use the measured distribution of voids as a way of inferring something of non-random distribution of galaxies and of clusters of galaxies.

John Maddox

Palaeontology

New fossil finds in old rocks

from S. M. Awramik

THE well-preserved stromatolites from 3,300–3,500-Myr-old rocks in South Africa described on page 489 of this issue¹ are not the first report of stromatolites from the Early Archaean, but these structures are by far the best preserved and morphologically convincing examples of such stromatolites that can be interpreted as microbial in origin. This discovery reinforces the notion^{2,3} that the Earth's earliest biota was much more advanced morphologically, and presumably metabolically, than has previously been thought.

Dealing with Archaean fossils is a tricky business. First, you have to convince the scientific community that the objects in question are really fossils; then demonstrate that the rocks in which the fossils were found are of the proposed Archaean age; and finally show that the fossils were coeval with the rocks. With almost every claim for the discovery of an Archaean fossil there are counterclaims proposing alternative non-biological origins, or that the rock or fossil in question is not of Archaean age⁴. Such prudence is justifiable because, in addition to extending the known record of life, Archaean fossils provide information on the diversity and relative abundance of early life, and lead to interpretations of metabolic characteristics and affinities of the microbes concerned. They also have significant bearing on evolutionary models that include the origins of life, its early diversification, phylogenetic relationships, rates of evolution and evolutionary events leading to the appearance of eukaryotes, animals and plants.

The general model for the early history of life used to be one of sluggishly evolving prokaryotes that took some 2,000 Myr to reach the eukaryotic grade and another 800 Myr before the animal level of organization evolved. The picture of life on the early Earth has become sharper in recent years as a result of important discoveries of fossils in the Early Archaean, 3,800–3,400 Myr ago. These discoveries have been made mainly in two regions: the Pilbara Block in Western Australia; and the Barberton Mountain Land in South Africa, areas that have excellent exposures of the oldest known relatively unmetamorphosed sedimentary rocks.

In rocks from the Swaziland Supergroup in the Barberton Mountain Land, solitary spheroidal microfossils, several micrometres across, that are interpreted as bacteria, have been detected in chert^{5,6}. Although these are currently regarded as 'dubiofossils', because there are doubts about their cellular microbial origin⁷, they

were headline news in the 1960s as they extended the first appearance of life by more than 1,000 Myr. Such small roundish microfossils were what most scientists would have predicted the Earth's oldest cells should have looked like. More convincing reports of spheroidal microfossils⁸ and the discovery of paired cells in the South African rocks soon followed⁹. The dyads suggested that these prokaryotes reproduced by binary fission and they made a stronger case for the biogenicity of the microfossils.

In 1980, two reports appeared^{10,11} on the discovery of stromatolites in the 3,400–3,500-Myr-old Warrawoona Group sediments from the Pilbara Block. Stromatolites present a very different picture when reconstructing the Early Archaean biosphere. They suggest more advanced microbial life. Stromatolites are conventionally viewed as biosedimentary structures produced by communities of photosensitive microbes that trap and bind sediment. At times, these microbes can generate layered domed and columnar structures centimetres to metres in size. Cyanobacteria (in particular filamentous cyanobacteria) and microbes that behave like cyanobacteria, are the dominant stromatolite builders today and were probably responsible for most fossil stromatolites. We do not know whether cyanobacteria existed in the Early Archaean but the stromatolites imply that proto-cyanobacteria, cyanobacteria-like microbes or some other prokaryote with stromatolite-building capabilities had already evolved the ability to live in shallow, agitated waters and cope with sediment and sunlight.

It is not surprising that the biological nature of the Warrawoona stromatolites has been questioned^{12,13}. Their morphology is simple, evaporites are associated with them and no microbial fossils were detected in their laminae. The growth of the evaporite minerals could have produced the simple doming. But this also occurs in modern evaporitic mat environments, where the shape of microbial mats (incipient stromatolites) is modified by the growth of evaporites.

Independent support for a microbial origin comes from the discovery of filamentous microbial fossils in Warrawoona chert³. Tubular fossil filaments suggest that the sheathed filamentous habit of prokaryotes, a highly successful stromatolite-building characteristic, has already evolved by the latter part of the Early Archaean. Unfortunately, these microfossils were not found in domed stromatolites. In summary, although the

evidence is strong, the jury is still out on the biogenicity of the Warrawoona stromatolites.

The focus for the oldest remains of life on Earth now returns to South Africa, to the Swaziland Supergroup, where the lowest part, the Onverwacht Group, is approximately the same age as the Warrawoona Group. Members of the research team at Louisiana State University that discovered stromatolites in the Warrawoona¹⁰ and microfossils in the Onverwacht¹⁴ now describe morphologically complex, biogenically convincing stromatolites from the Fig Tree Group, another part of the Swaziland Supergroup¹. These stromatolites are significant partly because their morphology is much more complex than the simple dome shapes found in the Warrawoona. The complex pseudocolumnar shape, good preservation, mode of occurrence, well-defined laminae and laminar microstructure are a combination of features that are difficult to explain by abiogenic processes. The closest abiogenic analogue would be geysirite. Furthermore, they are strikingly similar to other less ancient stromatolites, including early Proterozoic (2,500–1,600 Myr) biogenic stromatolites from the Gunflint Iron Formation and superficially similar to some modern non-marine cyanobacterial stromatolites. Combined with all the microbial fossil and organic geochemical data from the Early Archaean, the discovery adds to an impressive array of palaeobiological information on the biosphere of the early Earth.

It is unfortunate that stromatolites have played a relatively minor part in unravelling the early history of life. In many ways, the presence of stromatolites in such ancient rocks as the Swaziland Supergroup can reveal more about the degree of microbial evolution than microbial fossils preserved in chert, the common palaeobiological standard for the level of evolution in the Precambrian. Early Archaean stromatolites have interesting, if highly speculative, implications for the microorganisms that built them. First, as early as 3,500 Myr ago, microbes must have evolved that were benthic through-

1. Byerly, G.R., Lowe, D.R. & Walsh, M.M. *Nature* **319**, 489 (1986).
2. Knoll, A.H. *Paleobiology* **11**, 53 (1985).
3. Awramik, S.M., Schopf, J.W. & Walter, M.R. *Precambrian Res.* **20**, 357 (1983).
4. Cloud, P. & Morrison, K. *Precambrian Res.* **9**, 81 (1979).
5. Pflug, H.D. *Univ. Witwatersrand Econ. Geol. Res. Unit Info. Circ.* **28**, 1 (1966).
6. Schopf, J.W. & Barghoorn, E.S. *Science* **156**, 508 (1967).
7. Schopf, J.W. & Walter, M.R. in *Earth's Earliest Biosphere* (ed. Schopf, J.W.) 214 (Princeton University Press, 1983).
8. Muir, M.D. & Grant, P.R. in *The Early History of the Earth* (ed. Windley, B.F.) 595 (Wiley, London, 1976).
9. Knoll, A.H. & Barghoorn, E.S. *Science* **198**, 396 (1977).
10. Lowe, D.R. *Nature* **284**, 441 (1980).
11. Walter M.R., Buick, R. & Dunlop, J.S.R. *Nature* **284**, 443 (1980).
12. Buick, R., Dunlop, J.S.R. & Groves, D.I. *Alcheringa* **5**, 161 (1981).
13. Cloud, P. *Proc. 27th Int. Geol. Congr. (Moscow)* **5**, 173 (1984).
14. Walsh, M.M. & Lowe, D.R. *Nature* **314**, 530 (1985).

out most or all of their life cycle and were either fast-growing or motile — otherwise they could not have maintained their position at or near the sediment/fluid interface and would have been buried. A positive photoresponsive behaviour would enhance this ability, and presumably the microbes were photoautotrophic. Second, because the stromatolites formed in very shallow waters, they must have had some minimum resistance to high-energy solar radiation and finally, it is probable that some of the builders were filamentous forms with sheaths around the trichomes.

No longer can the evolutionary history of Precambrian life be viewed as slowly

changing microbial biota that gradually increased in diversity and complexity until the appearance of eukaryotes. The Swaziland stromatolites and the other Archaean fossil data indicate that by 3,300–3,500 Myr ago, diverse microbial life existed on Earth, some of which was organized into microbial ecosystems. Once life appeared on the early Earth, it diversified rapidly and in all probability explored and occupied most if not all habitable ecological zones. □

S.M. Awramik is in the Department of Geological Sciences, University of California, Santa Barbara, California 93106, USA.

Developmental genes

Mediators of cell communication?

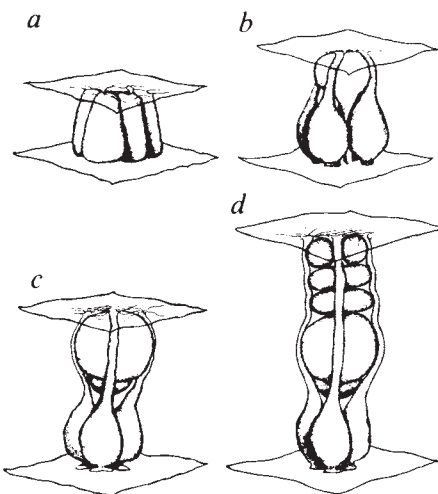
from Michael Akam

THE control of development operates at many different levels. In recent months this has become easy to forget, with the nucleus the focus of attention and the DNA-binding properties of the homeobox taking much of the limelight. Now the sequencing of two other 'controlling' genes, one from insects and the other from nematodes, has revealed homology with mammalian proteins of very different function. The *lin-12* gene of *Caenorhabditis*, which controls patterns of cell lineage¹, and the *Notch* gene of *Drosophila*, which plays a role in the early development of the nervous system², both encode proteins homologous with epidermal growth factor (EGF) and the structurally related family of mammalian proteins. Other members of this family are all membrane-bound or secreted proteins, suggesting that *Notch* and *lin-12* act at or beyond the cell membrane, perhaps to mediate cell communication.

The family of EGF-related proteins all share a characteristic cysteine-rich peptide of 40–50 amino acids. This sequence occurs nine times in the membrane-bound precursor of EGF; the terminal copy of the repeat, cleaved from the precursor, is the EGF peptide itself³. The *Notch* and *lin-12* genes also encode multiple copies of this peptide motif. The complete structure of the *lin-12* protein is not yet clear, but I. Greenwald¹ finds eleven copies of the EGF motif within the four exons that have been sequenced. *Notch* seems to encode an enormous protein of some 2,700 amino acids in a 10.5-kilobase-long transcript. K. Wharton and colleagues² have sequenced all but a few hundred bases of this transcript from overlapping complementary DNA clones. In the amino-terminal half of the derived protein they predict 36 tandem repeats of a 40-amino-acid motif that shares the conserved cysteine spacing of EGF, and shows limited homology at

other sites within the repeat. This region is followed by a putative transmembrane sequence of hydrophobic amino acids, followed by a large carboxy-terminal region which, except for a polyglutamine stretch, bears no close homology with other known proteins.

The structure of the *Notch* protein predicted by this sequence, which includes an extracellular EGF-like domain, resembles that of the membrane-bound protein in



Talking cells? The segregation of neuroblasts is an early step in the formation of the insect nervous system, in this case the grasshopper, and one which requires adjacent cells to communicate. The *Notch* mutations of *Drosophila* disrupt this process (see text). In insects, neuroblasts segregate from a sheet of apparently equivalent cells in the embryonic ectoderm (a). The first sign that a cell will become a neuroblast is the dorsal shift of its nucleus and cell body towards the interior of the embryo (b). Several cells in each region may initiate this process but only a single cell completes it. This cell will undergo the characteristic pattern of divisions which generate neurones (c,d). (Redrawn from ref. 9.)

the EGF family, for example, the low-density lipoprotein (LDL) receptor and the precursors of several growth hormones (EGF, tumour growth factor and vaccinia virus growth factor). But the EGF family also includes secreted proteins with diverse functions, including the blood-clotting factors urokinase and plasminogen activator. The role of the EGF domain within these secreted proteins is not known.

At first sight, the effects of *lin-12* and *Notch* mutations have little in common with one another or with the functions of the EGF protein family. The *lin-12* mutations cause losses and duplications of a range of structures in the animal, which result from alterations in the normally precise control of cell lineage during embryogenesis⁴. Recessive lethal *Notch* mutations (which cause notching in the adult wing when heterozygous) result in overgrowth of the nervous system in embryos⁵. There are, however, close parallels in the cellular phenomena underlying these different phenotypes. Both mutations alter the way in which equivalent cells choose between alternative developmental pathways. In both cases the decisions affected may depend on local interactions between neighbouring cells.

A functioning *lin-12* gene is needed for the normal development of several groups of cells during embryogenesis and maturation of *Caenorhabditis*. Cells of each affected group are characterized by analogous positions in the cell-lineage tree. In normal development, cells within each of these groups adopt two or more different fates (say A and B), each fate being defined by subsequent patterns of cell division and differentiation. When the *lin-12* gene does not function, all cells within a group are forced to adopt the same fate (fate A), or a subset of their normal fates. However, when mutation results in overexpression of the *lin-12* gene, all cells of the group are forced to adopt the alternative fate (B). Thus mutations that modulate *lin-12* activity switch the development of cells in a way that parallels the choices made during normal development. This suggested that the role of the *lin-12* gene in normal development was analogous to that of the homeotic selector genes in *Drosophila*⁶.

Several experiments where individual cells are killed show that the groups of cells affected by *lin-12* have the special properties of 'equivalence groups'⁷. In such a group, all cells seem to have the same intrinsic developmental potential; the fate adopted by each cell is specified by its location and by the interactions it makes with neighbours. One of the possible fates of a cell in an equivalence group is usually dominant; when a cell that will fulfill the dominant fate is killed, another member of the group alters its own behaviour to replace the dead cell.

Notch mutations seem to have a primary effect on neuroblast segregation. Recent descriptions of this process in insects^{8,9}, and of the effect that killing a single neuroblast has on the final pattern¹⁰, suggest that the determination of neuroblasts is in many ways analogous to the determination of cells within an equivalence group. In this case the neuroectoderm seems to comprise a sheet of equivalent cells, although the probability of neuroblasts forming is particularly high in certain regions of the sheet. In each of these regions, several cells may initiate the morphological changes leading to neuroblast formation (see figure), but one will come to dominate and inhibit its neighbours. In normal *Drosophila* embryos, about a quarter of the cells in the neuroectoderm become neuroblasts; the remainder make the larval epidermis. In *Notch* mutant embryos, it seems that the mechanism of domination and inhibition fails; all cells of the neuroectoderm become neuroblasts and no cells are left to form the epidermis.

The effects of *Notch* and of *lin-12* mutations could result from a failure at one of

the many levels in the hierarchy of control but, because of the homologies with EGF, it seems most likely that the defect lies in the ability of cells to sense or influence their environment. The gene product, at least in the case of *Notch*, is unlikely to be a hormone, because the effects of *Notch* mutations in genetic mosaics are expressed even in small regions of mutant tissue. The most attractive possibility is that *Notch* and *lin-12* are membrane proteins which transmit developmental signals between adjacent cells. □

1. Greenwald, I. *Cell* **43**, 583 (1985).
2. Wharton, K., Johansen, T.X. & Artavanis-Tsakonas, S. *Cell* **43**, 567 (1985).
3. Scott, J. *et al.* *Science* **221**, 236 (1983).
4. Greenwald, I., Sternberg, P.W. & Horvitz, R. *Cell* **34**, 435 (1983).
5. Lehman, R., Jimenez, F., Dietrich, W. & Campos-Ortega, J.A. *Wilhelm Roux Arch. dev. Biol.* **192**, 62 (1983).
6. Lawrence, P.A. *Nature* **306**, 643 (1983).
7. Kimble, J. *Dev. Biol.* **87**, 286 (1981).
8. Hartenstein V. & Campos-Ortega, J.A. *Wilhelm Roux Arch. dev. Biol.* **193**, 308–325 (1984).
9. Doe, C.Q. & Goodman, C.S. *Dev. Biol.* **111**, 93 (1985).
10. Doe, C.Q. & Goodman, C.S. *Dev. Biol.* **111**, 206 (1985).

Michael Akam is a Medical Research Council Senior Fellow in the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

Earth sciences

Sources of basic magmas

from R.N. Thompson

WHERE do basic magmas come from? Although more than 20 years have passed since it was generally agreed that the answer to this question is the Earth's mantle, there is still much debate and no unanimity as to the precise sites of basic magma genesis in the upper mantle. The source regions that are candidates for this role are the convecting asthenosphere and rigid lithosphere. The latter is absent immediately beneath mid-ocean ridges and can therefore be excluded as a source of erupted magmas at spreading centres. In contrast, both sources occurring in the older parts of ocean basins, near subduction zones and throughout the continents. The problem is that it is relatively easy to produce element and isotope analyses of basic rocks consistent with the genesis of their parental magmas by melting either a hypothetical lithosphere or asthenosphere mantle source, but it is extremely difficult to devise a real test between these alternatives. J.G. Fitton and H.M. Dunlop have recently devised such a test for a major group of basic magmas¹, consisting of basalts, with subordinate basanites and olivine tholeiites.

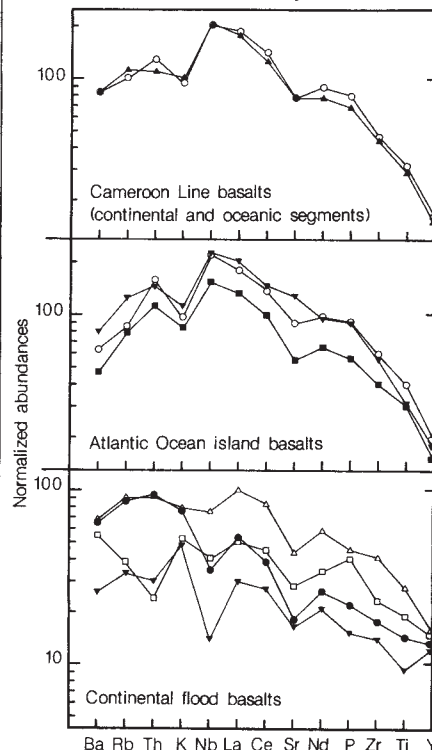
As a group these rocks are chemically indistinguishable from the basalts of ocean islands such as Hawaii and Iceland. Fitton and Dunlop's samples are from the Cameroon line, a 1,600-km chain of volcanoes. This immense igneous province runs

from the oceanic islands of Pagalu, Sao Tomé and Principe, in the Gulf of Guinea, across Cameroon to the Biu Plateau in north-east Nigeria, and straddles a continental boundary. The volcanoes of both the oceanic and the continental sectors of the line have been intermittently active throughout the past 25 Myr; most of them have erupted most recently within the past 1 Myr¹. Therefore, despite its suggestive appearance on the map, the line cannot represent the surface trace of an upwelling mantle plume beneath the African lithospheric plate, but rather it must be the locus of upward leakage from a long zone of basic magma generation in the upper mantle. The test presented by this zone is therefore a simple one. If lithospheric mantle does provide the source of magmas along the line, basalts from the oceanic and continental segments should be very different in their element and isotope composition because the sub-oceanic lithospheric in the Gulf of Guinea was formed at a spreading centre no more than about 125 Myr ago², whereas the sub-continental lithosphere beneath Cameroon is Precambrian. Fitton and Dunlop examined many trace elements and strontium isotopes, and showed that the average compositions of the basic magmas from the oceanic and continental segments of the Cameroon line are identical and therefore that the

magmas must have originated in the convecting upper mantle and cannot have interacted to any significant extent with lithosphere mantle¹.

The importance of this study lies in the ordinariness of the Cameroon-line lavas, which are typical ocean-island basalts, supporting earlier suggestions that these basalts contribute to continental magmatism³ and consistent with models that require the source mantle of ocean-island basalts to be continuous beneath the continents^{4–6}. Does this mean that oceanic and continental basalt magmatism (away from subduction zones) are the same phenomenon? Far from it: most continental basalts — especially those of the great lava floods — contain abundances of elements and isotopes from the oceanic compositional range. Explanations for these differences are usually either that the magmas originated from the same asthenospheric sources as oceanic basalts and reacted with continental crust on their way to the surface⁷, or that the magmas were formed by melting of subcontinental lithospheric mantle which had developed a distinctive composition during prolonged isolation from the convecting asthenosphere⁸.

Mafic magmas are hundreds of degrees hotter than the liquids of the most fusible parts of continental crusts, which makes it difficult to understand why evidence for



Abundance of a range of elements in basalts from the Cameroon line and elsewhere, normalized to facilitate synoptic comparisons. The Atlantic Ocean island samples are from Tenerife, Madeira and St Helena, the continental flood basalt samples from the Columbia River Plateau and Snake River Plain, USA, the Parana Basin, Brazil, and Karoo, South Africa.

crustal contamination of the liquids during magma upwelling is not found more often. H.E. Huppert and R.S.J. Sparks⁹ have recently provided one solution to this problem by investigating the fluid dynamics of subterranean magma migration in regions of volcanism associated with lithospheric extension, such as the Cameroon line, where the magmas probably rise in dykes. They calculate that the rheology of the upwelling liquids varies from laminar to turbulent within the geologically relevant range, turbulent behaviour occurring mainly at the high flow rates associated with the large eruptions that characterize flood basalts. Huppert and Sparks point out that a basic magma moving in a laminar manner through a conduit forms a chilled skin along the contact which tends to prevent incorporation of potentially fusible wall-rock into the fluid. In contrast, a turbulent magma will actively erode any patches or layers of fusible material along the margins of its conduit.

This argument does not apply only to magma–crust interaction. The Mg-rich, low-viscosity ultramafic magmas, which are identified by phase-equilibria studies to be the parents of flood basalts^{10–11}, are also likely to behave turbulently as they rise through the subcontinental lithosphere. They are hot enough to melt and incorporate selectively the relatively fusible ‘metasomatic’ veins and patches that characterize subcontinental lithospheric mantle⁸. It therefore seems possible that fluid dynamics holds the key both to the eruption of unmistakably oceanic asthenosphere-source basic magmas along the continental segment of the Cameroon line (small magma batches traverse the lithosphere by laminar flow, with minimal interaction), and to the contrasting compositions of almost all flood basalts (large and fast-moving magma batches rise turbulently through the subcontinental lithosphere and incorporate sufficient of its most fusible constituents to emerge with compositions suggesting an ancient lithospheric source). □

1. Fitton, J.G. & Dunlop, H.M. *Earth planet. Sci. Lett.* **72**, 23 (1985).
2. Le Pichon, X. & Huchon, P. *Earth planet. Sci. Lett.* **67**, 123 (1984).
3. Allègre, C.J., Dupré, B., Lambret, B. & Richard, P. *Earth planet. Sci. Lett.* **52**, 85 (1981).
4. McKenzie, D.P. & O’Nions, R.K. *Nature* **301**, 229 (1983).
5. Thompson, R.N., Morrison, M.A., Hendry, G.L. & Parry, S.J. *Phil. Trans. R. Soc. A* **310**, 549 (1984).
6. Allègre, C.J. & Turcotte, D.L. *Geophys. Res. Lett.* **12**, 207 (1985).
7. Carlson, R.W., Lugmair, G.W. & MacDougall, J.D. *Geochim. cosmochim. Acta* **45**, 2483 (1981).
8. Hawkesworth, C.J., Rogers, N.W., Van Calsteren, P.W.C. & Menzies, M.A. *Nature* **311**, 331 (1984).
9. Huppert, H.E. & Sparks, R.S.J. *Earth planet. Sci. Lett.* **74**, 371 (1985).
10. Cox, K.G. *J. Petrol.* **21**, 629 (1980).
11. Thompson, R.N., Morrison, M.A., Dickin, A.P. & Hendry, G.L. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C.J. & Norry, M.J.) 158 (Shiva, Nantwich, 1983).

R.N. Thompson is at the Department of Geology, Imperial College, London SW72BP, UK.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

The sand-bubbler crab *Scopimera intermedia*, showing the unique air-breathing structures, called gas windows, on the legs. In this species the gas windows are red. Each leg membrane (there are two per leg) is very thin (0.6 µm) and covers a large area (up to 25 mm²). A complex capillary-like network of blood vessels lies beneath. See p. 493 for a complete description. (Photograph by David P. Maitland.)

Human evolution

How small was the bottleneck?

from J. S. Jones and S. Rouhani

THE human fossil record is no exception to the general rule that the main lesson to be learned from palaeontology is that evolution always takes place somewhere else. Even the best fossil beds preserve evidence of about one skeleton per million¹, and this, combined with problems of dating, means that the pattern of origin of modern *Homo sapiens* remained obscure until recently. It is not surprising that palaeontologists have at different times hailed Europe, Africa, Asia and even the Americas as the birthplace of modern man². Mankind, however, retains a more complete record of history in its genes, and a paper on page 495 of this issue³ provides new evidence that the origin of modern man lies in Africa.

In their study, Wainscoat *et al.*³ examine the patterns of relatedness of a small segment of nuclear DNA from five closely linked restriction sites in the β-globin gene cluster from Europeans, Indians, Asians and Africans. A total of 32 haplotypes (combinations of alternative sites) can be generated from these 5 variants; but in the event only 14 are found among the 601 chromosomes studied, and only 5 of these are common. There is a striking geographical pattern in their distribution. Three β-globin haplotypes are widespread in all non-African populations (and are rare in Africa) whereas the African population is characterized by a high frequency of two haplotypes, one of

which is rare and the other absent from the rest of the world. For this segment of DNA, mankind can be divided into an African and a non-African lineage.

The replacement of Neanderthal man by anatomically modern *H. sapiens* in a few thousand years in Europe (about 30,000 years ago) and fossils similar to modern humans have been found in Asian beds which may be as much as 50,000 years old^{2,4}. It is sometimes suggested that the relatively sudden appearance of widespread populations of modern man means that evolution of the species occurred independently in different parts of the world². This does not agree with modern theories of speciation: the probability of a single transition between the adaptive peaks that represent species is low, and the chance of a simultaneous peak shift in an evolving lineage is small^{5,6}. The replacement of Neanderthals and other archaic humans by modern man in most parts of the world is much more likely to have resulted from the extinction of the older population by new hominids spreading from a single centre of origin. Expanding human populations did spread remarkably quickly. Most palaeontologists agree that the Americas were devoid of mankind until about 12,000 years ago, but once man had entered the empty New World he reached the southern tip of the continent in less than 1,000 years⁷. The wave of advance of early farmers in Europe, measured by the

appearance of their artefacts and by the gradients in gene frequency which persist in modern populations⁸, occurred at about 1 km a year across Europe from 6,000 BC — an advance sufficiently fast to pass from Greece to England in two millennia.

Wainscoat *et al.*³ suggest that the pattern of distribution of β -globin haplotypes is consistent with the rapid spread of modern man from a centre of origin in Africa, and that the spread of the species to the rest of the world led to a drastic shift in haplotype frequencies because of a population bottleneck and sampling drift in the emigrant population. New fossil evidence strongly supports the view that man's origin was indeed Africa. Anatomically modern human fossils from 100,000 years ago have been found at the mouth of the Klasies River in South Africa and modern man was widespread in Africa by 50,000 years BP (ref. 9), at a time when Europe was populated by Neanderthals. Studies of blood group and enzyme alleles in Africa and of DNA sequences associated with the sickle-cell mutants suggest that African populations are genetically more subdivided than are those from other parts of the world^{10,11}. This again suggests a longer history of evolution *in situ* in Africa than anywhere else.

When did modern man leave Africa and how small a population emigrated? Although the evidence from fossils is sparse, their distribution suggests that the first emigrants had left their continent of birth by 50,000 BP and were widespread by 30,000 BP. The social structure of modern African hunter-gatherers does predispose them to rapid genetic changes resulting from sampling drift¹². However, the global extent of β -globin divergence has at first sight some startling demographic implications because the hunter-gatherers who migrated from Africa. Europe and Asia have rather similar haplotype frequencies. Hence, the emigrants must have undergone the major change in haplotype frequency in the interval between leaving Africa and dispersing throughout the rest of the world. Assuming — and this is little more than an informed guess — that this interval was 20,000 years, population-genetics theory tells us that the mean effective size of the ancestral population for all non-Africans throughout this period must have been 600 individuals; or alternatively that the bottleneck was 6 individuals for 200 years, or even a single couple for 60 years. (The expected time for the loss of a neutral gene present in the population at frequency p is $E(T) = -4N \ln p / (1-p)$, where N is the population size¹³. We assume a generation interval of 20 years and that the 4 common haplotypes were present at equal frequencies in the ancestral African population.) If this is the case, much of mankind was an endangered species during an important part of its evolution.

Population bottlenecks of this severity in modern human populations can have untoward genetic consequences. Africaners returned to the land of their African ancestors about 300 years ago, and most of them can trace their families to one of a few Dutch founders. One partner of a couple who married in 1688 carried an allele for the recessive disease porphyria variegata. This allele is now carried by many of their descendants and is hundreds of times more common among white South Africans than elsewhere¹⁴. If the distribution of β -globin haplotypes is typical of those in the human genome, founding populations of non-African man may have suffered a considerable genetic load because of inbreeding. There are some puzzling inconsistencies when the global patterns of these haplotypes are compared with those of other human polymorphisms. These can give widely different estimates of the size of the human population during its flight from Africa.

Sequence data for mitochondrial DNA (which is maternally inherited) suggest that African populations are ancestral to others¹⁵. Casanova *et al.* have found analogous differences between African and European peoples in a Y chromosome-linked DNA polymorphism (which is paternally transmitted and hence may be useful in studying gene flow through males¹⁶). Calculations of the size of the bottleneck that can explain the relatively lower mitochondrial diversity of non-African populations compared with Africans suggest a population of 4,000 individuals for 20,000 years, or 40 individuals for two centuries. Mitochondrial divergence therefore implies that mankind underwent less of a bottleneck on emigration from Africa than the estimate derived from β -globins. Global patterns of blood group and enzyme diversity are on the average much more uniform geographically than are those of either β -globin or mitochondrial DNA sequences: the mean difference between Africa and Europe is not much greater than that which exists between two Africa tribal groups or two European countries¹⁷. For these genes there is no need to postulate a bottleneck during the emergence from Africa very different from those which have occurred in man's immediate past.

Why should different genes produce such different estimates of the size of the founding non-African population? Mitochondria have peculiarities in their mode of transmission and mutation rate which limit their value as indicators of evolutionary relatedness¹⁸, and natural selection on some enzymes and blood groups may obscure their demographic history¹⁹. Most important, statistical problems mean that estimates of the time and pattern of divergence and of the size of evolving populations based on individual DNA sequences have large systematic

errors^{20,21}. Such estimates depend greatly on the number of unlinked loci surveyed, so that single short sequences of closely linked markers may be misleading indicators of population history. These uncertainties can be seen in a recent study of 43 sequence variants in a 2,400 base pair region of the alcohol dehydrogenase locus of the fruitfly *Drosophila melanogaster*²². Although the genetic information available for *Drosophila* is much greater than is that for human β -globin, it makes it possible to estimate the date of origin of the enzyme polymorphism only to within a range of 0.6–3.5 Myr BP.

These statistical problems mean that the phylogeny of individual genes may be a misleading indicator of the phylogeny of the populations from which they come. Quite often two alleles will diverge before the date of splitting of an ancestral population²³. This is already clear from studies of human populations: nobody claims that the common high frequency of the B blood group (an allele that is absent from many human populations) which is found in Russians and chimpanzees is a reliable indicator of a shared evolutionary history! The same caution must be observed when reconstructing human evolution from other individual genes or DNA sequences.

It is thus possible that future sequence studies of the human genome will cast a different light on population bottlenecks during the spread of modern man. However, unlike students of the human fossil record, molecular biologists know that the information they need to reconstruct human evolution is easily available and merely awaits their attention. □

1. Cronin, J.E. *et al.* *Nature* **292**, 113 (1981).
2. Howell, F.C. in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) xiii (Liss, New York, 1984).
3. Wainscoat, J.S. *et al.* *Nature* **319**, 495 (1986).
4. Stringer, C.B. *et al.* in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) 51 (Liss, New York, 1984).
5. Barton, N. & Charlesworth, B. *A. Rev. Ecol. Syst.* **15**, 133 (1984).
6. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7641 (1985).
7. Owen, R.C. in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) 517 (Liss, New York, 1984).
8. Ammerman, A.J. & Cavalli-Sforza, L.L. *The Neolithic Transition and the Genetics of Populations in Europe* (Princeton University Press, 1984).
9. Brauer, G. in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) 327 (Liss, New York, 1984).
10. Nei, M. & Roychoudhury, A.K. *Evol. Biol.* **14**, 1 (1982).
11. Pagnier, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 1771 (1984).
12. Harpending, H. & Wandsnider, L. in *Curr. Dev. Anthr. Genet.* **2**, 29 (1982).
13. Kimura, M. & Ohta, T. *Genetics* **61**, 893 (1969).
14. Dean, G. *The Porphyrias: a Study of Inheritance and Environment* 2nd edn (Pitman Medical, London, 1971).
15. Johnson, M.J. *et al.* *J. molec. Evol.* **19**, 255 (1983).
16. Casanova, M. *et al.* *Science* **230**, 1403 (1985).
17. Latter, B.D.H. *Am. Nat.* **116**, 220 (1980).
18. Wilson, A.C. *et al.* *Biol. J. Linn. Soc., Lond.* **26**, 375 (1985).
19. Piazza, A.A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 2638 (1981).
20. Tajima, F. *Genetics* **105**, 437 (1983).
21. Takahata, N. & Nei, M. *Genetics* **110**, 323 (1985).
22. Stephens, J.C. & Nei, M. *J. molec. Evol.* **22**, 289 (1985).
23. Hudson, R.R. *Evolution* **37**, 203 (1983).

J.S. Jones and S. Rouhani are in the Department of Genetics and Biometry, University College London, London NW1 2HE, UK.

X-ray astronomy

New results from Exosat

from A.C. Fabian

THE European X-ray astronomy satellite Exosat has been in orbit for more than two years. Its orbital period of almost 100 h allows a long uninterrupted view of X-ray sources to an extent that was impossible with earlier satellites in near-Earth orbits under the radiation belts, and has led to the many discoveries discussed at a recent meeting*. X-ray astronomers are used to each successive satellite probing deeper in a discovery space that has coordinate axes in spatial, spectral and temporal resolution, together with sky and time coverage. Exosat excels in this last respect, with further advances made possible by its good response to very soft X-rays and the high spectral resolution to photons at the energy of the iron line.

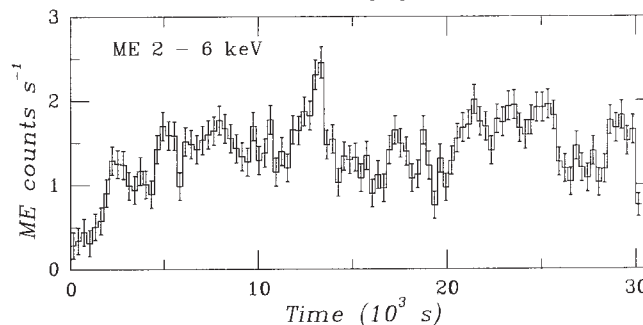
The recently discovered narrow galactic ridge (Warwick, R. *et al.* *Nature* **317**, 218; 1985) has been studied in more detail (R. Warwick, University of Leicester). Some bright spots on the ridge are supernova remnants, but these cannot explain all of the ridge. The spectrum of the 'diffuse' emission is consistent with that of gas at $kT=8$ keV, as found independently by observers using the Japanese satellite Tenma. There are many point sources that are probably X-ray binaries, some of which are new. One such transient X-ray source, EXO0748-676, discovered while the satellite was slewing between targets, was discussed by A. Parmar and M. Gottwald (European Space Operations Centre, ESOC). The source is a member of the class of 'dipping' X-ray sources which display recurrent shallow dips in their light curves. The source also shows X-ray bursts, the recurrence time of which correlates strikingly with the persistent source luminosity in a manner opposite to that suggested by current theories. Several bursts were followed by a second burst within 20 min or less.

Exosat monitoring of the well-known jet emitter SS 433 was reported by G. Stewart (University of Leicester). A strong iron-emission line at about 7 keV was Doppler-shifted in phase with the 164-day precession cycle of the jets. As only one line is observed, that caused by the approaching jet, instead of two lines as seen at optical wavelengths, Stewart concludes that the iron emission originates from close to the central object. The other jet is presumed to be obscured by an accretion disk.

Most excitement was centred on the discovery of quasi-periodic oscillations (QPOs) in several low-mass X-ray bi-

naries. These were first reported by M. van der Klis and colleagues (*Nature* **316**, 225; 1985) from the source GX5-1. Power spectra of the light curve show a low-Q peak that shifts according to the apparent luminosity. There are now about seven sources that display this phenomenon, which in many respects resembles the behaviour of some cataclysmic variables observed at optical wavelengths. S. Ilovaisky (University of Besançon) pointed out that QPOs are also observed in the candidate black hole GX339-4. They are found optically when the source is either very bright or faint, but not when it is in any intermediate state. G. Hasinger (Max Planck Institute for Extraterrestrial Physics) summarized the X-ray QPOs from CygX-2 which are at about 40 Hz, a much

Light (X-ray) curves of NGC4051, showing rapid movements (K. Pounds).



higher frequency than any yet observed optically.

It is not known why QPOs were not observed by previous X-ray satellites, although they were detected in 1982 during some flat-topped bursts of the rapid burster with the satellite Hakucho. J. Swank (NASA Goddard Space Flight Center) has analysed Einstein satellite data and found no QPOs. Her limit for CygX-2 is about 3 per cent, whereas reports from Exosat give 7 per cent. The long-duration Exosat observations produce clean power spectra that must contribute to the detection of a weak, fluctuating signal. W. Friedhorsky (Los Alamos) found QPOs in the brightest continuous X-ray source in the sky, SCOX-1, which has been extensively studied; the QPOs are only seen when the source is not flaring. His data and the new results of van der Klis show that the source behaviour in the QPO frequency-intensity plane is no simple correlation but may show complete open loops.

Stella (ESOC) reported on QPOs in GX17+2 and the rapid burster, where the frequency behaviour is again complex. The limits on regular pulsations in all these objects is often better than 1 per cent

and makes demands on some of the models for QPOs (Lamb, F.K. *et al.* *Nature* **317**, 681; 1985) that rely on magnetic effects, as stressed by J. van Paradijs (University of Amsterdam). To have a sufficiently strong magnetic field to cause QPOs from the difference between the magnetosphere and disk rotation rates, he argues that the neutron stars must be young ($<5 \times 10^7$ yr). The stars then need evolved companions so that a high rate of mass transfer could have pushed their progenitor white dwarf over the Chandrasekhar limit. We are likely to hear a lot more about QPOs as they are one of the few observational handles that we have on low-mass X-ray binaries.

The more detailed studies of X-ray bursts made possible by Exosat have revealed unusual events. A. Tennant (University of Cambridge) reported on a train of bursts observed from CirX-1, a source previously thought to be a candidate black hole. As all bursts yet seen appear to originate from neutron stars, we must reassign CirX-1. The bursts observed by

Tennant have a short recurrence time of about 40 min and occur half-way between the characteristic 16.6-day outbursts.

Short recurrence times were also found in MXB1705-44 by M. Sztajno (Max Planck Institute for Extraterrestrial Physics). There, small bursts occurred only a few minutes after several larger bursts. The generally accepted model for the Type I X-ray bursts, which show a rapid rise followed by a cooling decay, invokes a thermonuclear flash of the accreted helium. This may also trigger off rapid hydrogen burning, which at lower temperatures is restricted to a slow rate by β -decays in the CNO cycle. The spectrum of the emission is usually fitted to that of a black body from which source size and luminosity can be estimated as long as the distance is known.

This approach has often led to problems with source radii smaller than that of theoretical neutron stars computed using current equations of state and bursts that seem to exceed the Eddington limit. The problems can in part be resolved by a correct model atmosphere calculation for the spectrum rather than a simple black body. A. Foster (University of Cambridge) has fitted the spectra to bursts from MXB1728-34 and obtained luminosities at sub-Eddington limits for a sensible radius and acceptable source distance.

Exosat has also been used to monitor active galactic nuclei (AGN). The extra-soft response of the low-energy telescopes over that of the Einstein observatory has

*Cambridge Exosat discussion meeting, Institute of Astronomy, Cambridge, UK; 5-6 November 1985.

paid off with the discovery of soft X-ray excesses in the spectra of several AGN. G. Branduardi-Raymont (Mullard Space Science Laboratory), in spectra of NGC5548, and K. Pounds (University of Leicester), in a survey of 27 galaxies, find such excesses; 18 of the 27 AGN show a typical power-law spectrum with a photon slope of 1.71 ± 0.11 (see figure). Line-of-sight absorption above that expected from within our Galaxy has been detected in 10 objects. None of this X-ray absorption has yet been observed to vary. Several AGN sources vary on timescales of hours and days, the most rapid variability being reserved for the lowest luminosity sources such as NGC4051. A. Lawrence (Queen Mary College, London) saw this source vary by factors of around two on a timescale of 1,000 s and P. Barr (ESOC) saw similar amplitude changes in the nucleus of M81 in 10 min. His work suggests that there is a good correlation between the fastest time-resolved variation timescale of an AGN and its luminosity.

The rate of change of the luminosity of Seyfert galaxies indicates that electron-positron pairs are common. A. Treves

(Milan) has found an extreme change in the BL Lac object PKS 2155-304 where the X-ray luminosity increased from 1.5×10^{45} to 10^{46} erg s⁻¹ in about 3×10^3 s. Unless relativistic effects are present, the mass-to-energy conversion efficiency exceeds 50 per cent.

The high observing efficiency of Exosat, which is not affected by the South Atlantic anomaly and only rarely by Earth occultation, has produced an enormous quantity of data that would have taken about five years to collect from a near-Earth orbit. So far, the main results have arisen from the exploitation of light curves and power spectra. The spectral results will follow when the detailed calibrations are made. P. Giommi (ESOC) has compiled a database of serendipitous sources that has barely been investigated. And yet more new discoveries can be expected from the remaining year of Exosat's life (assuming successful perigee boosts next January) both from the satellite itself and from the databank. □

A.C. Fabian is at the Institute of Astronomy, The Observatories, Madingley Road, Cambridge CB3 0HA, UK.

Neurobiology

Microcircuits to see by

from Gordon M. Shepherd

IT HAS been known for 20 years, from work in the retina¹ and olfactory bulb², that neuronal dendrites can send synaptic signals as well as receive them and can take part in microcircuits for local information processing. In virtually every vertebrate species studied it has been established that all but one of the main cell types in the retina have axons and/or dendrites that are both presynaptic and postsynaptic (*a* in the figure). The conspicuous exception has been the ganglion cell, the long-axon projection neurone that sends the retinal output to the lateral geniculate nucleus, the next station in the visual pathway. The report on p.495 of this issue by Sakai, Naka and Dowling³ brings the ganglion cell in out of the cold and shows that its dendrites can be presynaptic as well as postsynaptic.

Working with the catfish retina, Sakai *et al.* labelled ganglion cells with horseradish peroxidase (HRP) so that ganglion cell bodies and dendrites containing HRP reaction product could be identified in electron micrographs. Using serial sections the authors could distinguish between proximal and distal dendrites, and were able to characterize the types of synaptic connections. Intracellular recordings of the ganglion cell responses to light stimuli before HRP injection permitted correlations between physiological properties, dendritic branching pattern

and synaptic organization.

The analysis confirmed the previously accepted pattern of synaptic connections on ganglion cells, as shown on the right in *a*. They also found the connections shown on the left of *a*; distal dendrites of large-field ganglion cells making synapses on vesicle-containing processes (amacrine or ganglion cell dendrites) and ribbon-containing processes (bipolar cell axon terminals). The bipolar and amacrine cell terminals receiving these synapses were grouped in clusters around the distal dendrites. But these patterns of synaptic connections were found only on large-field ganglion cells with varying properties (on-off, on-centre, off-centre) — by contrast, no presynaptic dendrites or distal terminal clusters were seen around small-field ganglion cells.

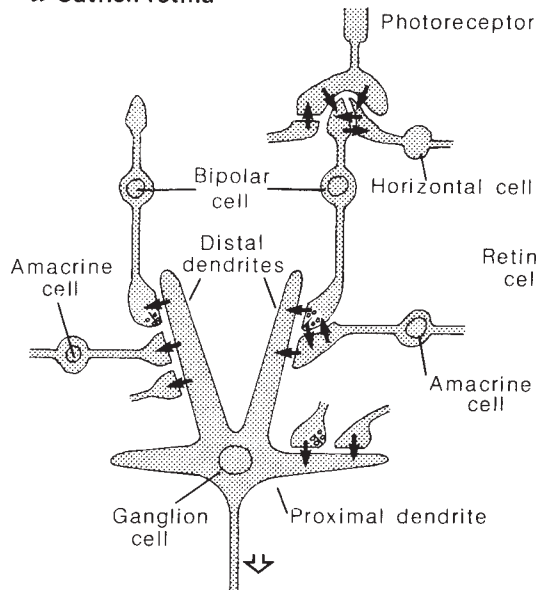
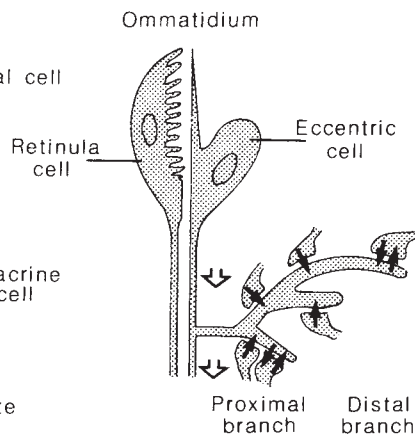
An interesting parallel to this work is a recent study by Fahrenbach⁴ of the synaptic organization of the compound eye of *Limulus*, the horseshoe crab. The classical work of Hartline and collaborators showed that the eccentric cells in *Limulus* eye are involved in lateral inhibitory interactions with each other, taking place through synaptic connections made by eccentric cell neurites in the neural plexus that lies just below the eye. Using serial sections through the plexus Gur *et al.*⁵ demonstrated both single and reciprocal synapses that could mediate the lateral

inhibition; they believed that these connections were made at or near the site of impulse initiation in the axon.

Fahrenbach has examined a series of more than 1,000 sections with loss of only 1 section, covering an area of 22 ommatidia and extending from the eccentric cell bodies to the base of the neural plexus. From this he was able to reconstruct the complete synaptic connectivity of all the neurites from 54 eccentric cells that branched within one small knot (microglomerulus) of neuropile in the plexus. He found no synapses on the axon itself, but collaterals close to the point where they branched from their axons of origin had a high synaptic density and a predominance of input synapses; terminals of branches far from the axon had a low synaptic density and a predominance of output synapses. Up to 50 per cent of the synapses were arranged in reciprocal pairs (*b* in the figure). Many terminals appeared to be grouped in clusters.

The single-cell analyses in the retina are technically demanding and the exhaustive reconstruction in *Limulus* even more so: are such herculean efforts worth it? The answer lies in the insights gained about the relationships between structure and function and the principles they suggest. In vertebrates, the retinal ganglion cell can now be classed with the olfactory mitral cell² as a long-axon neurone with dendrites that are both presynaptic and postsynaptic. The findings of Sakai *et al.* add to the evidence that synaptic triads (tight patterns of connections between afferent terminals, relay neurones and interneurons) involving presynaptic dendrites are a dominant theme of the synaptic organization of the visual pathway at every level up to the cortex. For example, this type of organization has been further documented in the lateral geniculate nucleus by Hamos *et al.*⁶ In arthropods, several neurones with long axons connecting intersegmentally⁷ or to muscles⁸ have branches that are both presynaptic and postsynaptic; some patterns of synaptic connections of the branches resemble those seen in the eccentric cell. Integration of synaptic inputs and outputs within

1. Dowling, J.E. & Boycott, B.B. *Proc. R. Soc. B* **166**, 80 (1966).
2. Rall, W., Shepherd, G.M., Reese, T.S. & Brightman, M.W. *Exp. Neurol.* **14**, 44 (1966).
3. Sakai, M.S., Naka, K. I. & Dowling, J.E. *Nature* **319**, 495 (1986).
4. Fahrenbach, W.H. *Proc. R. Soc. B* **225**, 219 (1985).
5. Gur, M., Purple, R.L. & Whitehead, R. *J. gen. Physiol.* **59**, 285 (1972).
6. Hamos, J.E., van Horn, S.C., Raczkowski, D., Ulrich, D.J. & Sherman, S.M. *Nature* **317**, 618 (1985).
7. Watson, A.H.D. & Burrows, M. *J. comp. Neurol.* **205**, 383 (1982).
8. Selverston, A.I., Russell, D.F., Miller, J.P. & King, D.G. *Prog. Neurobiol.* **7**, 215 (1976).
9. Ralston, H.J. *Nature* **230**, 585 (1971).
10. Koch, C., Poggio, T. & Torre, V. *Phil. Trans. R. Soc. B* **298**, 227 (1982).
11. Peters, B.H., Altman, J.S. & Tyrer, N.M. *J. comp. Neurol.* **233**, 269 (1985).
12. Nishimura, Y. & Rakic, P. *J. comp. Neurol.* **241**, 420 (1985).

a Catfish retina**b Limulus eye**

Aspects of the organization of presynaptic dendrites and microcircuits in vertebrate and invertebrate eyes. The sites and polarities of individual chemical synapses are indicated by solid arrows, the direction of spike conduction by open-headed arrows. *a*, Catfish retina: the main types of retinal cells are shown, with some of the main synaptic connections. The connections shown on the right of the ganglion cell are the well-established inputs; on the left, the newly described ganglion cell output synapses on to bipolar cell processes and amacrine or ganglion cell processes³, found on distal dendrites of ganglion cells (see text). *b*, Compound eye of *Limulus*: schematic representation of an ommatidium containing retinular cells and an eccentric cell. The eccentric cell is analogous to the ganglion cell in having a long axon that projects to higher visual centres. The synaptic connections made by a branch in the neural plexus are indicated. There is a predominance of input synapses on proximal branches and a relatively larger proportion of output synapses distally⁴. The significance of this distribution is discussed in the text.

neuronal branches thus seems to be a common mechanism in arthropods.

The ganglion cell is known to lack recurrent axon collaterals and has been believed to function entirely as a follower cell, driven by its inputs from bipolar and amacrine cells. The new evidence³ indicates that at least some of the ganglion cells can take part in shaping their own response properties through the microcircuits formed by their presynaptic den-

drites. Sakai *et al.* point out that the synaptic clusters around distal dendrites can function as relatively independent sites for local signal processing, which also seems to apply to the clusters in *Limulus*. Individual or clustered synapses functioning as integrative units are becoming familiar in many systems⁹⁻¹¹. The degree to which the relative independence of such units is influenced by passive membrane properties and by active dendritic sites is under

investigation in several laboratories.

How are dendritic input and output synapses laid down during ontogeny? This is the subject of recent work by Nishimura and Rakic¹² in the inner plexiform layer of the primate. In work in progress, they find a sequence in which single ganglion cell synapses appear first, followed by single amacrine – amacrine synapses, bipolar ribbon synapses on ganglion cells and finally, bipolar ribbon synapses on amacrine cells. This sequence appears to be genetically determined, because it is not correlated with birth or exposure to light. Ganglion cell dendrites have not been observed to be presynaptic in this study but it will be interesting to use the HRP labelling technique of Sakai *et al.* to pursue this point.

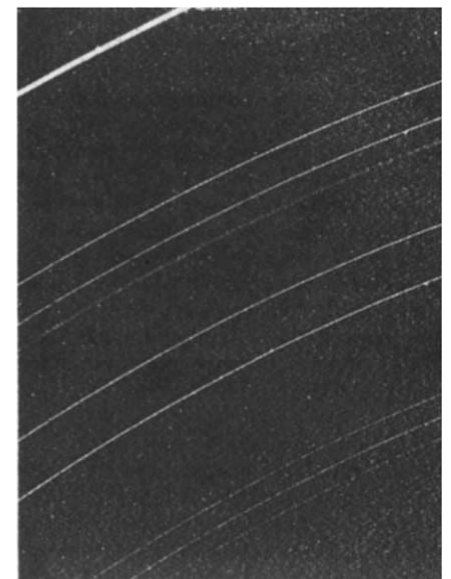
In *Limulus*, Fahrenbach speculates that the relatively uniform distribution of output synapses on proximal and distal branches would be determined by the genome during synaptogenesis; by contrast, the distribution of input synapses, heavily skewed toward the proximal branches, might be determined by a gradient of impulse activity spreading decrementally from the axon into the arborization. In this scheme of activity-generated synaptogenesis, "any branch would in effect provide the signal to generate its own inhibitory input, graded with distance from the parent axon"¹⁴. This would account for the gradation of inhibitory interactions.

How gene expression is synchronized in two or more sets of neurones to bring about the coordinated assembly of synaptic circuits, shaped by concurrent or subsequent functional activity, is one of the most challenging problems in neurobiology. Microcircuits make excellent test systems for pursuing this problem. □

Gordon M. Shepherd is Professor in the Section of Neuroanatomy, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

Rings and a moon of the planet Uranus

On the right are the ten rings of Uranus, photographed from a distance of 1,120,000 km on 23 January 1986. This is among the first of the pictures produced by Voyager 2. The newly discovered tenth ring is barely visible about midway between the bright, outermost epsilon ring and the next ring down, called delta. The tenth ring orbits Uranus at a radius of about 50,000 km. On the left is a full-disk view of the moon Titania, taken early in the morning of 24 January 1986 from a distance of 500,000 km. The many circular depressions are probably impact craters and the bright linear troughs right of the picture are probably canyons. The troughs break the crust in two directions, indicating that the icy satellite has a dynamic, active interior. Photographs courtesy of NASA/Jet Propulsion Laboratory.



Large computers key to aggregation phenomena

SIR—We thank you for devoting a part of a *News and Views* editorial¹ to one of our recent papers². Our paper describes a reversible process of particle-cluster aggregation-disaggregation in which particles at the periphery of a cluster may escape and come back again after undergoing a random walk in space. We agree with your argument that this does not truly represent aggregation at equilibrium, but the motivation behind our work was somewhat different. We aimed to describe more exactly some kind of dynamic equilibrium or steady state in which a flux of incoming diffusing particles strictly compensates a flux of outgoing diffusing particles.

In this spirit, our work was and attempt to include restructuring processes in the existing models of aggregation (other attempts are by Meakin and Jullien³ as you quoted, but also P. Meakin⁴). This was a first step before studying cluster-cluster aggregation-disaggregation processes, where bonds can be annealed randomly all along the diffusion-sticking mechanism. This last model was recently studied by Kolb⁵.

We agree that our main result (our clusters have apparently the same fractal dimension as animals) is quite intricate. When describing our model to several colleagues, we got a large range of guesses and speculations; they predicted either DLA, or Eden, or animals and so on. This is why our paper included a detailed discussion explaining that it was none of those. However, since we found a fractal dimension close to the one for animals, we insisted on the comparison with this at-equilibrium model.

Interestingly, however, cluster-cluster aggregation-disaggregation, as demonstrated by Kolb, also yields the same fractal dimension. Moreover, in the past⁶, Stauffer used a process very similar to ours (but without any diffusion) in which particles are randomly released from the perimeter and then put back randomly on it (this is a "chemical" version of our aggregation-disaggregation model). Stauffer claimed that his objects were in the same universality class as animals.

It is very surprising that these three aggregation-disaggregation models all share the same fractal dimension, since different rules might be expected to induce different statistics. The cluster statistics for small clusters are clearly not the same as for large clusters, so why do they peak to an average configuration in the thermodynamic limit of very large clusters?

To conclude, we think that as long as there is no analytical progress in these fields, large size computer simulations will

be the *only* tool enabling us to understand and classify the models. Moreover, they can define the basis for a search for analytical solutions. In the case of aggregation phenomena the two basic models (particle-cluster and cluster-cluster aggregation) are now well characterized and they must be modified and extended to better apply to experiments (compare for example the recent attempt to take into account cluster polarizability⁷, to explain the experiment by Hurd and Schaefer⁸). It is however essential that in the course of these improvements, the models do not lose their high degree of simplicity and generality.

R. BOTET
R. JULLIEN

*Université de Paris-Sud,
Centre d'Orsay,
Laboratoire de Physique des Solides,
91405 Orsay, Paris, France*

1. Maddox, J. *Nature* **318**, 229 (1985).
2. Botet, R. & Jullien, R. *Phys. Rev. Lett.* **55**, 1943–1946 (1985).
3. Meakin, P. & Jullien, R. *J. Physique* **46**, 1543–1552 (1985).
4. Meakin, P. *J. chem. Phys.* **83**, 3645–3649 (1985).
5. Kolb, M. *J. Phys. A*, (in the Press).
6. Stauffer, D. *Phys. Rep.* **54**, 1–74 (1979).
7. Jullien, R. *Phys. Rev. Lett.* **55**, 1692–1697 (1985).
8. Hurd, A. & Schaefer, D. *Phys. Rev. Lett.* **54**, 1043–1046 (1985).

Minimum energy configurations

SIR—Berezin¹ has recently pointed out that the minimum energy configuration for a system of m identical point charges interacting via the Coulomb potential $1/r$ and confined to a circular region in a plane has all the charges on the circumference only if $m \leq 11$. For $m > 11$ some of the charges are in the interior. (Specifically, Berezin showed that the energy of configuration B ($m-1$ charges uniformly distributed around the circumference plus one charge at the centre) is lower than that of configuration A (m charges uniformly distributed around the circumference) for $m > 11$. We have extended Berezin's work and have shown that at $m = 17$ it is advantageous to have two particles in the interior. Note also that for $m \rightarrow \infty$ the problem becomes that of the conducting disk.) He conjectured that a similar result might obtain in three dimensions. Unfortunately one can show, using the fact that the Coulomb potential is harmonic in three dimensions², that the minimum energy configuration for a system of m identical point charges confined to a region of space has all the charges on the surface of the region, independent of the number of particles m . If, however, we change Berezin's conjecture slightly, the problem again becomes interesting.

We ask: what is the minimum energy configuration for a system of m identical particles interacting via a non-harmonic potential $1/r^n$, $n > 1$, and confined to a

spherical volume? (This is why Berezin's two dimensional system has charge in the interior — the potential $1/r$ is not harmonic in two dimensions. If, as MacGowan³ suggests, one uses the potential $\ln(r)$, harmonic in two dimensions, the charge all lies on the circumference.) Clearly, for $m = 2, 3, 4$, the particles will lie on the surface of the sphere independent of n . For higher m the problem becomes more difficult.

Melnik *et al.*⁴ have studied in detail the problem of the minimum energy configuration for systems of particles constrained to the surface of a sphere. Calculations based on their configurations lead us to conjecture that the minimum energy configuration has all the particles on the surface for $m \leq 12$, independent of the power n of the potential. For $m = 13$ and higher, however, the situation is different. This can be seen most easily in the short range limit $n \rightarrow \infty$. In this limit the energy is dominated by nearest neighbour interactions. At $m = 13$ the nearest neighbour distance becomes less than 1 (in units of the sphere radius) and it becomes advantageous to remove one particle from the surface and place it at the centre of the sphere (leading to a minimum energy configuration of icosahedron plus one at the centre). We have shown that this situation persists if the power of the interaction n is as small as approximately 6.3.

For $n < 6.3$ the minimum energy configuration has all the particles on the surface in the configuration described by Melnik *et al.*⁴. It seems likely that a similar situation will persist for still larger numbers of particles: for low powers of the interaction all the particles will be on the surface; for higher powers some particles will be in the interior. A complete mapping of the " m - n phase plane", will, however, involve much detailed study.

M.G. CALKIN
D. KIANG
D.A. TINDALL

*Department of Physics,
Dalhousie University,
Halifax, Nova Scotia,
Canada B3H 3J5*

1. Berezin, A.A. *Nature* **315**, 104 (1985).
2. Kellogg, O.D. *Foundations of Potential Theory* 223 (Springer, Berlin, 1929).
3. MacGowan, D. *Nature* **315**, 635 (1985).
4. Melnik, T.W., Knop, O. & Smith, W.R. *Can. J. Chem.* **55**, 1745–1761 (1977).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a rather technical character which are not provoked by articles or letters previously published (where the Matters Arising section remains appropriate).

All change for the RI

Roy Porter

The Royal Institution: An Informal History. By Gwendy Caroe.
John Murray: 1985. Pp.180. £13.95.

WHEN insiders write institutional history, all too often they serve up an official version for public consumption; conversely, outsiders generally lack a feel for the place. Who better, then, for avoiding both of these shortcomings in tracing the story of the Royal Institution than the late Gwendy Caroe? As the daughter of Sir William Bragg, and the sister of Sir Lawrence, she was steeped in the atmosphere of Albemarle Street. From the time of her mother's early death, Gwendy was hostess at her father's weekly scientific dinner parties there (she tells a hair-raising story against herself, of when she flipped a glass of claret over H.G. Wells's shirt just before he was due to lecture). Yet, never having been on the establishment, she can look at its august, if somewhat chequered, history with proper detachment.

As all who have read her biography of her father will know, Gwendy Caroe wrote with rare grace, lucidity and a vivid turn of phrase. In this new book those assets are again matched with an eye for telling detail and a wise gift for getting to the heart of people — Faraday, whose science was animated by profound religious humility and a dread of worldly clamour; Tyndall who sought society because he was essentially lonely; Davy who craved it, because his vanity couldn't bear not being fêted. The chapter on Sir James Dewar in particular brings a neglected yet fascinating character to life.

It would be wrong to assume that just because the book is subtitled *An Informal History*, and is not weighed down by footnotes and graphs, it is merely a piece of light reading. Far from it. Not only is this the only account we have of the RI from its foundation in 1799 to the present, but it is a penetrating and original interpretation at that, the fruit of much delving in the archives. Not least, Gwendy Caroe shows how the RI's fortunes are a story of permanent change.

Amongst institutions, the RI has always been something of a curiosity, being neither fish, flesh nor fowl; neither simply a learned society, nor a specialist research agency, nor chiefly a centre of scientific training, nor an organ of state. It has carried its essentially protean nature right from its first foundation, when, as Gwendy Caroe stresses, one of its key aspirations was to bring together the worlds of science and fashion. As science has marched on, and fashions have come and gone, so the RI has changed with them.

The initial idea was that, through promoting applied science, the RI could help the rich do their bit for the poor. Its early patrons — the benevolent but ambitious Count Rumford, the evangelical Thomas Bernard, the autocratic president of the Royal Society, Sir Joseph Banks — and the grandees whose subscriptions made them Managers saw the Institution as a philanthropic laboratory. It would develop technologies to better the condition of lower orders in that age of French Revolutionary turmoil: improved kitchen utensils, fuel economies, dietary innovations, simpler soap-boiling and the like. Meantime, lecture courses teaching trades to artisans would foster independence and self-improvement. And all of these good causes would be financed via scientific lectures delivered to High Society.

The fashionable lectures thrive — Davy in particular dazzled the duchesses — but the philanthropy was soon dropped. Instead, the RI took on an important function as a centre of scientific analysis for aiding farmers and manufacturers. At the Managers' behest, Davy and then Faraday devoted themselves to long years of practical investigation into tanning, dyeing, soil science, miners' safety lamps, optical glasses and scores of other utilitarian subjects, to show how science could be the willing servant of an industrializing society. Yet eventually the professors dug in their heels. First Faraday, and then more particularly John Tyndall, insisted

that the best return for the Institution would come through the pursuit of what Faraday called "the advancement of science for its own sake". Faraday's great electro-magnetical discoveries, Tyndall's explorations of heat, pressure and gaseous radiation, and Sir James Dewar's achievements in the liquefaction of gases were some of the fruits.

Yet the move in the latter part of Victoria's reign towards pure science, allowing the professors to get on with their own private research interests and consultancy work, left the Institution as such in the doldrums. It ceased to lead society, state and even fashion. Its subscriptions dropped off. By the time of the First World War, even its splendid new research laboratory (the Davy-Faraday), so handsomely financed by Ludwig Mond, was gathering dust.

New directions were needed. They came from the Braggs. In turn, in the 1920s and the 1950s, they brought fresh collective purpose. Sir William pioneered group research at the RI, developing his own X-ray crystallography research school. Father and son both revitalized the "club" functions and the public lecturing programme, pioneered contacts with the BBC and put School Science Lectures on the footing that does so much to keep the RI in the public eye today. Not least, under their guidance, the old shoe-string private experimentation gave way to research more adequately funded from government sources. Over the past 20 years with Sir George Porter at the helm, the RI has continued to branch out in new directions. What has been demonstrated is that though the RI cannot pretend to fulfil all the functions of the Royal Society, university science institutes, the research councils and the like, there continues to be a need for a popular forum for science.



Dazzling the Duchesses — a Gillray cartoon of a lecture at the Royal Institution in 1802, "Scientific Researches — New Discoveries in PNEUMATICKS!". Davy holds the bellows.

which the RI is uniquely equipped to provide.

It has always been easy to be somewhat snooty about the RI. Fellows of the Royal Society have often found it vulgar (William Buckland spoke sneeringly of "a sort of Albemarle Street dilettante meeting"); and today, a school of radical historians decries it for having been a "tool of capitalism". Yet, as Gwendy Caroe shows, the RI has played a crucial role in English science: it created an open forum for science, amateur and professional, in contrast to the Royal Society's oligarchy and subsequent elitism; it supported an industrial laboratory in days when manufacturers were still poo-pooing science; it offered a refuge for religious Dissenters such as Faraday long before they could gain admission to English universities, and a livelihood for promising young talent — Davy was 23 and Faraday 21 when they came to the RI — before science had evolved a regular professional career structure; and not least, it provided a bridge between the "two cultures", and between experts and the public. Science and fashion have truly met at the RI. For an institution so often capable of meeting the challenge of the new, the future surely holds exciting prospects. □

Roy Porter is at The Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK.

**MAJOR
NEW JOURNAL**

**Alimentary
Pharmacology
& Therapeutics**

Editors:
Dr. R.E. Pounder, Royal
Free Hospital, London
Professor M.J.S. Langman,
University Hospital,
Nottingham

To be published
bi-monthly from
January
1987

Further details can be obtained from
Allen Stevens,
Journal Marketing Manager,
Blackwell Scientific
Publications,
Oxford OX2 0EL,
UK

Reader Service No.4

Looking beyond Bohr

Philip Pearle

Quantum Fluctuations. By Edward Nelson. Princeton University Press:1985. Pp.146. Hbk \$32, £21.50; pbk \$12.95, £9.40.

IT WAS Niels Bohr who gave quantum theory the closest of possible shaves with Occam's razor, in declaring that comprehension of quantum phenomena with classical ideas was impossible, and that we must never never never try to understand what goes on in nature in between measurements. How well it speaks for the vigour of physics, just over 100 years after the birth of Bohr*, that a book should appear attempting all that he told us not to do.

In Nelson's picture of nature, classical point particles moving under classical forces are jostled by a fluctuating hypothetical "background field". This requires a statistical description of the motion. Such a description inevitably entails two equations: one deals with the ensemble of particles, ensuring that none of them get lost ("conservation of probability"); the other is an equation of motion for each individual particle. These two equations are shown to be precisely equivalent to the imaginary and real parts, respectively, of the Schrödinger equation.

Nelson, more than anyone else who has worked with such ideas, has made the equation of motion plausible. The difficulty is that plausibility often means analogy, and there is none. A speedily moving jostled particle (such as a mote of dust battered by air molecules) will not maintain that velocity, not even on the average the way a quantum particle will, but will slow down and come to thermal equilibrium with its environment. Such a classical motion is called a *dissipative* (because the particle loses energy) *Markovian* (meaning that the present, and not past history, determines the future) *diffusion* (referring to the dispersal of particles with initially identical position and velocity).

Lacking an analogy, Nelson has invented "stochastic mechanics", a non-dissipative (conservative) Markovian diffusion process, in which particles do not lose their energy on the average. The first half of the book is a masterly instruction in diffusive motion on manifolds (curved spaces), with particular attention to elucidation of the time-reversal-symmetric description. It culminates in a very natural variational principle (largely due to Guerra and co-workers), the minimization of

* A celebration of Bohr and his achievements, *Niels Bohr: A Centenary Volume*, edited by A.P. French and P.J. Kennedy, will be reviewed in *Nature* of 20 March.

the expectation value of the action, which gives the desired equation of motion. This is equivalent to, but more elegant than, an earlier (1966) formulation of Nelson's, in which he wrote down Newton's second law of motion but with the acceleration of the particle defined in a particular time-symmetric way.

Next come examples of things that quantum theory never talks about. We learn that particles never cross the nodes of wavefunctions. We see the electron jiggling in hydrogen, drifting towards the Bohr radius in the ground state, and rotating about the nucleus on the average in a non-zero angular momentum eigenstate. There are explanations of half-integer spin, of statistics, and of the quantum theory expression for the momentum probability distribution.

It is in the discussions of locality and of two-slit interference that challenges to stochastic mechanics emerge. Nelson shows that, in this theory, disturbance of one of two widely separated particles instantly disturbs the diffusion of the other particle, which is a non-local behaviour, and poignantly writes: "I have loved and nurtured Markovian stochastic mechanics for 17 years, and it is painful to abandon it. But its whole point was to construct a physically realistic picture of microprocesses, and a theory that violates locality is untenable". (He then argues that non-Markovian motion may not have this defect.) It is interesting that upon encountering this difficulty, David Bohm, the author of a similar theory (differing from Nelson's in that the equation of motion is deterministic), has embraced it, delighting in the principle that everything is connected to everything else.

I believe it is two-slit interference that best illustrates a difficulty many physicists have with this theory. Nelson clearly shows how particles go through either one slit or the other and then diffuse to form the interference pattern ("... this is a misnomer: nothing is interfering with anything"). But such diffusion is effectively determined by the wavefunction of quantum theory, and it is hard to see how a "background field" can give rise to the effects of the wavefunction. Examples of a background field causing conservative diffusion are needed.

Nelson himself appreciates the point: he is probably his own theory's severest critic. He writes: "Stochastic mechanics is quantum mechanics made difficult". Physicists may well ask why they should consider the complications that stochastic mechanics brings to quantum theory. The answer is that it is worth the difficulty of opening our minds, to see what may be revealed if we attempt to peer past the curtain erected by Bohr. □

Philip Pearle is Chairman of the Department of Physics, Hamilton College, Clinton, New York 13323, USA.

What is theoretical biology?

Lee A. Segel

Theoretical Biology and Complexity: Three Essays on the Natural Philosophy of Complex Systems. Edited by Robert Rosen. Academic: 1985. Pp.210. \$57, £48.

IN HIS preface, Robert Rosen states that the three scientists whose essays make up this book

... each began with a conviction that contemporary physics already contained the necessary universals with which to cope with the phenomena of life and that therefore only a clever rearrangement and redeployment of these universals would suffice to bring them to bear effectively on biology. ... we each came separately, and with great reluctance, to admit the possibility that this conviction might not be true and hence that a true theory of the organism required new physics and new epistemology. And again separately, we realized that the measurement process, which lies at the very heart of every mode of system description, provides perhaps the only safe and fundamental point of departure for building a comprehensive theory, not only of organisms, but of natural systems in general.

I believe that most workers in theoretical biology are indifferent to, or disagree with, the main points just cited. But as I have no brief to speak for others, let me give my own views.

It is inappropriate to try to build comprehensive theories at this early stage in the modern study of biology. Rather, it is better to tackle particular phenomena and to try to generalize from the results. Although there is probably need to redress a present-day over-confidence in the comprehensiveness of conclusions that can be reached by molecular biology, the remedy is not to concentrate upon epistemology.

Certainly, an eventual aim should be to understand "the organism"; that is, answer the question "What is life?". But there are other, equally important and more approachable questions to be asked, such as "How does a baby develop from a fertilized egg?" or "How does the brain process, store and retrieve information?". Of course, new physics is needed for this enterprise—for example, even recent extensions of thermodynamics apply only to situations that are close to equilibrium. The major province of present-day theoretical biologists is to extend old concepts and introduce new ones in phenomenological fashion, at appropriately selected levels in the hierarchy of biological organization.

To my mind, saying that the measurement process is central to biology is to relapse into the sin of adopting too uncritically the ideas of physics. An example of

a more successful point of departure is the stress on computation in the study of vision by D. Marr and his associates.

The authors are of course entitled to have their say. In an extension of classical formulations of mechanics and thermodynamics, I.W. Richardson discusses the interaction between particles and the consequent observable manifestations thereof. (There is no explicit mention of biology in his essay.) A.H. Louie tries to deal with problems of measurement, interaction and representation using category theory, a relatively recent branch of abstract algebra. (More than one hundred special symbols are listed at the beginning of the essay, signalling the fact that it will be comprehensible to few apart from mathematicians.) Here, too, there is little specific contact with biology.

Rosen's article does contain quite a bit about biology and is relatively accessible. But for two reasons I doubt the need for his urgings to go beyond the evolving state description that has been so successful in particle physics—such descriptions have in fact proved valuable in large areas of biology (for example in biochemical

theories that exploit the law of mass action), and alternative formalisms have frequently been employed (for example graph theory in ecology).

The authors' point of view is best summed up by Dr Louie, who states that "I have always been more interested in 'conceptual' biomathematics than in looking at specific models of a biological process. In other words, my interest is oriented more toward the *logic* of mathematical biology". Those who share this attitude and who have a good mathematical background will find food for thought in this book. Those who are more pragmatically inclined should peruse the *Journal of Theoretical Biology*, *Biophysical Journal*, the "Models and Hypotheses" section of *Differentiation* or some of the more than 60 volumes in Springer-Verlag's series *Lecture Notes in Biomathematics*. Included in such publications, albeit embedded in inescapable chaff, are many specific kernels of understanding that theorists have brought to biology. □

Lee A. Segel is a Professor in the Department of Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel.

On the track of an epidemic

Richard S. Tedder

Understanding AIDS: A Comprehensive Guide. Edited by Victor Gong. Rutgers University Press/Cambridge University Press:1985. Pp.240. Hbk \$20, £17.50; pbk \$9.95.

AIDS: The Acquired Immune Deficiency Syndrome. By Victor G. Daniels. MTP Press:1985. Pp.155. £9.95, \$16.25.

AIDS: Etiology, Diagnosis, Treatment and Prevention. By Vincent DeVita, Jr., Samuel Hellman and Steven Rosenberg. Lippincott/Harper & Row:1985. Pp.352. \$38, £41.95.

TO CALL Acquired Immune Deficiency Syndrome the infectious disease of the 1980s may unhappily be wishful thinking—although this decade has seen the emergence of AIDS, it is most unlikely to see its departure. Initially thought to be limited to homosexual men, there is no doubt that the causative virus, HTLV III/LAV, is becoming widely established in the human population. It is especially unfortunate that the homosexual man bears the brunt of infection, since this has engendered both a moralistic and a narrow view of the disease. Moralistic, in the sense that people who do not have homosexual lovers can underline their disapproval by invoking Divine Retribution; narrow, in that a belief that the virus is confined to gay men encourages

complacency about public health.

It is the speed with which HTLV III/LAV has established itself as an *endemic* infection within the human race that should be the main cause for concern, and not merely the increasing numbers of patients with AIDS. Taken alone, the figure for cases of AIDS is a gross underestimation of the prevalence of infection and the true size of the problem for the future. Early in any epidemic the initial cases of disease will be those with a short incubation and will underrepresent the clinical attack rate. There is now clear evidence that the virus may be transmitted by heterosexual intercourse and is not only passed on by homosexual activity. This is particularly true of the African epidemic which is itself of recent origin; moreover, women may transmit the disease to the next generation, either *in utero* or possibly through breast milk. The human race is facing a microbial threat which, rightly or wrongly, must call into question current attitudes to all sexual practices. It is surely not sufficiently comforting to remark that AIDS is essentially a disease of certain limited segments of the population when we know that infected persons may be both infectious and themselves subject to illness of diverse forms for an entire lifetime.

Set against this background, any book called *Understanding AIDS: A Comprehensive Guide* is in for a certain amount of scepticism. Where knowledge is accruing so fast, any book on the subject will be out of date as soon as it is published. However in this instance it may never have been "in date". Although the

editorial preface was written in February 1985, the book itself appears to have been written largely before HTLV III/LAV was acknowledged as the causative agent. The various contributions are generally adequate but read in a very disjointed manner. Though understandable this is annoying — it is as if the galley proofs were returned with a request to mention HTLV III/LAV “since it might be relevant”. Perhaps the publishers would have been wise to ask for a full revision. There are no illustrations or figures and the bibliography is rather sparse.

The two other books are different from each other and from the previous one. They serve different functions and will be of value to different readers. *AIDS* by Victor Daniels is a concise book, “in English”, and in the first 100 pages provides an up-to-date compilation of current knowledge about AIDS and its relation to infection by HTLV III/LAV. The remaining 50 pages include a glossary, and appendices on blood donors, immunology, viruses, definitions and a useful list of contact addresses of support groups in Britain. This is not a scientific textbook, but it will be invaluable to any non-specialist or lay person and of particular use to those involved in counselling of infected and affected patients.

The last book comes with a picture of the HTLV III retrovirus (with all credits) on the cover, so its pedigree cannot be called into question. Scientifically it is the most robust of the three, and is a well-rounded collection of review articles, each fully referenced and well illustrated. In particular it is a pleasure to see good colour prints of both the clinical lesions and the histology of Kaposi's sarcoma including patients with classical non-epidemic disease — an illness only rarely encountered outside Africa. One feels mild irritation in certain chapters at the parochially American view point — for example reference to European work, such as that of the Pasteur group, is minimal. The chapter on diagnostic tests is disappointing, with Western blotting still firmly placed as the gold standard, and the molecular virology is sketchy. However these are small quibbles and it remains a book to be recommended for workers in the field or the technically inclined non-specialist.

What none of these books does is to speculate on the origin of HTLV III/LAV and question the established dogma that Africa is the source of the infection; nor does any of them give an indication of possible future courses of the epidemic. On these topics, perhaps, too many reputations are at stake. Only time will tell. □

Richard S. Tedder is Senior Lecturer and Head of the Virology Section in the Department of Medical Microbiology, School of Pathology, Middlesex Hospital Medical School, Riding House Street, London W1P 7PN, UK.

Current prospects

Daniel D. Eley

Molecular Semiconductors: Photoelectric Properties and Solar Cells. By J. Simon and J.-J. André. *Springer-Verlag: 1985. Pp.288. DM174, \$72.50.*

IN THIS book, consideration of molecular semiconductors is restricted to organic semiconductors, apart from brief sections on two types of inorganic conductors (tetracyanoplatinates and sulphur-nitride polymers). The last comprehensive account of the subject as a whole appeared in 1974, and even then a list of 2,579 references was included. It seems logical, therefore, that Simon and André should now concentrate on one section of the field currently of great interest. There is at present little indication that an organic photovoltaic cell will be able to compete with amorphous silicon in solar applications, but all the more reason for the publication of a book such as this to help stimulate research in the area.

After a short introduction to basic notions in solid state physics, the authors set out in a second longer chapter their approach to photoelectric phenomena in molecular photocells. They characterize their treatment as involving localized-states; it might also be called quasi-chemical. A fairly full account of Schottky junctions is included, stating further that the various characteristics of these contacts are quite easily obtained from alternating current studies. I would dispute this, since hopping mechanisms within the bulk of the organic can also give rise to alternating current dispersions and so complicate the overall picture. Here, too,

the authors could have said with advantage something about photocarrier lifetime and photoconduction gain, which must be important for a successful solar energy device.

The next chapter deals with metallo-phthalocyanines and Chapter 4 with polyacetylene. These are followed by a chapter giving brief accounts of some ten or so other systems, including the inorganics already referred to, and superconductors. The list of 1,191 references is a useful guide to the relevant literature.

In a one-page conclusion, the authors rightly point out that the low mobilities found in organic semiconductors limit their possible applications. One might add that not enough efforts have so far been made to measure mobilities in organics, and to establish mobility-structure relationships. The authors further conclude that “no physical or theoretical considerations limit the possibility of obtaining molecular semiconductors”. Unfortunately chemical stability problems, such as those that bedevil linear polyene and especially iodine-doped systems, do limit useful applications. In this respect it is sensible that the longest chapter is concerned with phthalocyanines, renowned for their stability (although they are not completely free from such problems, and also the difficulties associated with purification).

This is a useful and well-written book, of reasonable length. It should attract attention from the solid-state community in general, but it will be of particular interest to chemists working on the conduction properties of molecular semiconductors. □

Daniel D. Eley is Emeritus Professor of Physical Chemistry at Nottingham University, University Park, Nottingham NG7 2RD, UK.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

The largest wind-tunnel in the world under construction at Ames Research Center, California, in 1943. The picture is reproduced from Searching the Horizon, by Elizabeth A. Muenger, a history of Ames from 1940 to 1976. The book is published by NASA as SP 4304 and is available from the US Government Printing Office, price pbk \$13.

Optical synchrotron emission in the southern lobe of 3C33

K. Meisenheimer* & H.-J. Röser

Max-Planck-Institut für Astronomie, Königstuhl, D-6900 Heidelberg 1, FRG

Charge coupled device polarimetry at a wavelength of 660 nm has detected a highly-polarized optical source, which is coincident with the radio hotspot at the leading edge of the southern lobe of 3C33. The optical polarization data perfectly match the high-frequency Very Large Array observations, thus providing the first direct evidence for optical synchrotron radiation from a classic double radio source. The extended structure of the emission region requires a huge particle accelerator boosting electrons to highly relativistic energies (up to 100 GeV) over a region several kiloparsecs in extent.

THE double radio source 3C33 consists of two extended lobes ~ 4 arc min apart from each other¹. Coincident with the central D galaxy ($z = 0.0592$ (ref. 2)) there is a compact radio core¹. Thus, 3C33 is a typical example of the powerful classic edge-brightened double radio sources, most commonly found in surveys at low frequencies (for example, 3C-, 4C-, Bologna-survey; extended radio sources are reviewed in ref. 3). At a resolution level of a few arc seconds high surface-brightness radio hotspots were detected in many extended radio lobes. For the radio hotspot located at the outer edge of the southern lobe of 3C33 (refs 4–6) a close association with a faint diffuse optical object has been reported⁷.

The nature of such associations is unclear as there is little information about the optical emission processes involved. Furthermore, significant mismatches between the suggested optical counterparts and the radio maxima are frequently encountered⁸, strongly enhancing the probability of chance coincidences with faint field objects.

The detection of optical synchrotron radiation from extended radio lobes would have special significance, since the path length of the highly relativistic electrons responsible for this radiation is short (≤ 1 arc s for redshifts $z \geq 0.05$ and typical magnetic field strengths of ~ 10 nT).

Thus, optical synchrotron emission would directly map the regions where particle acceleration takes place. In addition, the overall nonthermal spectrum and the energy distribution of the relativistic electrons would be fixed over a much wider range, leading to more reliable estimates of the total emitted power and to details of the acceleration mechanism involved^{3,9,10}.

The most striking characteristic of synchrotron radiation is its high linear polarization. The detection of a highly polarized optical object in the vicinity of a radio lobe would, therefore, establish direct evidence for optical synchrotron emission. Here we present our first results of CCD polarimetry in the fields of extended radio lobes. The southern lobe of 3C33 was chosen to start with because of its bright and compact hotspot and the reported optical emission.

Observations

The optical field around the southern lobe of 3C33 was observed on 14, 15 and 16 November 1984 by the 2.2-m telescope on La Silla (Chile). Polarization analysis of all objects in the field was made possible by replacing one of the filter wheels in front of the direct imaging CCD camera (thinned RCA chip, 45 e^- read-out-noise) by a rotatable plane parallel double calcite plate. It splits the light from each object in the field into two (perpendicularly) polarized images which are recorded simultaneously on the CCD detector (see Fig. 1, details of the method and the analysing procedure are given in ref. 11). The observations were

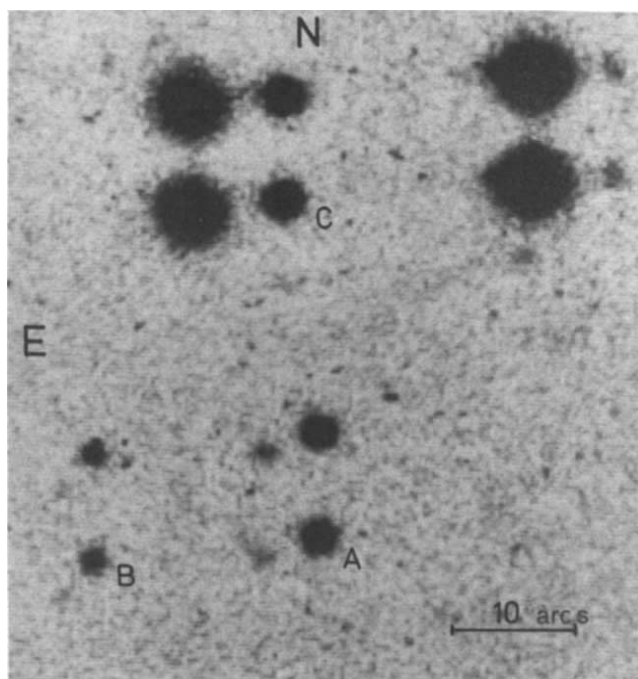


Fig. 1 Optical field around 3C33 south. This frame is a superposition of three CCD exposures (20 min each) taken through the double calcite plate in a red bandpass ($\lambda = 660$ nm). Each object in the field is split into two polarized images, separated in north-south direction. Southern images of reference stars are labelled A, B and C, respectively. The optical counterpart of the hotspot located 5 arc s south-east from star A is highly polarized as is obvious from the large brightness difference between the two images.

made in a red bandpass ($\lambda_0 = 660$ nm, $\Delta\lambda = 150$ nm (full width at half maximum)).

After the first 30-min exposure, a highly polarized optical object near the hotspot of 3C33 south was immediately conspicuous (see Fig. 1). Two subsequent exposures, taken with different position angles of the calcite plate, confirmed the detection. Three additional sets of frames with different orientations of the calcite plate were obtained in the following nights. Each set consists of three 20-min exposures with a fixed orientation of the calcite plate but slightly different pointing of the telescope.

For the detailed analysis of optical morphology and polarization of the hotspot counterpart each single CCD frame had to be corrected for sensitivity variations over the CCD detector and interference fringes. The appropriate flat field frames were

* Present address: Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK.

generated by averaging all exposures of the same night which were taken through the same filter and at similar zenith distances. To avoid distortions of the flat field frames due to the astronomical objects present on each exposure, we employed a κ - σ test¹² among the (normalized) values of each pixel in different frames, which were exposed with different telescope pointing. Discrepant pixel values are subsequently ignored in calculating the average flat field frames.

After the flat field correction, all exposures taken with the same orientation of the calcite plate are collated. On the resulting fully corrected frames the remaining background variations do not exceed $\pm 0.3\%$ on scales ≥ 2 arc s.

Optical counterpart of 3C33 south

The relative position of the highly polarized optical object and the radio hotspot in the southern lobe of 3C33 is determined with respect to the stars labelled A and C in Fig. 1, for which Rudnick *et al.*⁶ measured coordinates. The radio hotspot position is taken from the 15-GHz map with 0.85 arc s resolution published by Dreher⁵. We find a positional offset (optical-radio) $\Delta\alpha = -0.03 \pm 0.02$ s and $\Delta\delta = +0.1 \pm 0.3$ arc s, so in what follows we assume a common radio and optical position of the hotspot maximum located at $\alpha_{1950} = 01^{\text{h}} 06^{\text{m}} 12.10$ s, $\delta_{1950} = +13^{\circ} 02' 27.5''$.

As the extended structure of the optical counterpart is obvious, even on a single exposure, we superimpose all 18 images present on the three sets of CCD frames to obtain an image with optimum signal-to-noise ratio (Fig. 2a). The optical brightness distribution reproduces the four uppermost contours on the 5-GHz VLA (Very Large Array) map⁶ very well (see Fig. 2b). The backward tail of the hotspot pointing towards the central galaxy as well as the roughly perpendicular wing towards the north-west are clearly detected at $\lambda = 660$ nm. The main difference with respect to the radio map is the longer backward tail observed on our red image. A careful comparison with the blue 200-in plate⁷ indicates that this backward tail is of similar length in blue light, while hardly any blue radiation is detected from the hotspot itself (positional reference is based on the stars A and B, see Fig. 1).

Optical polarization

Since the optical counterpart of 3C33 south is clearly resolved, we attempt to obtain a polarization map. For this purpose, we select 11 different positions on the optical image at which we evaluate the optical polarization separately. The polarization analysis basically requires the comparison of the integrated counts within corresponding apertures on the two polarized

images and is done in the same way as for the jet of 3C273 described elsewhere¹¹.

The optical polarization in 3C33 south, evaluated for an effective beamwidth of 2.0 arc s, is given in Table 1, columns 4 and 5. In addition, the polarization analysis was carried out for all other objects in the field and for about 50 (empty) positions around 3C33 south, producing a very reliable estimate of the actual measuring errors involved. For comparison, the apparent polarization values as obtained for two positions located in the extremely steep brightness gradient of star A are also included in Table 1 (labelled A-1 and A-2). They demonstrate the perfect matching of corresponding apertures and exclude any severe distortion of the polarization measurements in 3C33 south due to this nearby 20 mag star A.

The optical polarization map (Fig. 3) shows that the polarization position angle of the hotspot coincides with the direction of the backward tail, as observed at 6 and 2 cm (ref. 6). Crossing the hotspot and the adjacent wing from the south-east to the north-west, a rotation of the polarization angle by $\sim 40^\circ$ is observed at both optical and radio frequencies (see Table 1). The mean optical polarization of the hotspot (average over positions 1 to 5) is $P = (29.2 \pm 2.4)\%$ at a position angle of the *E* vector $PA = 54^\circ \pm 4^\circ$. In the backward tail, just inwards of the hotspot (position 10) the orientation of the *E* vectors seems to flip abruptly by about 90° . Although we failed to reach a 3σ significance level in the amount of polarization at these faint brightness levels (≥ 24 mag/arc s², that is $\leq 2.5\%$ of the total background surface brightness which is doubled through use of the calcite plate), this 90° flip is in accordance with the magnetic field orientation inferred from the radio maps (compare with positions 10 and 11 in Table 1).

For a quantitative comparison with our optical results we measure the radio polarization at the 11 positions on enlarged copies of the maps published by Rudnick *et al.*⁶ (6-cm data are given in columns 6–8 of Table 1). Using the six positions near the radio hotspot for which the optical polarization exceeds 3σ we find a mean rotation of the polarization angle $\Delta\chi$ (6 cm-opt.) = $\{PA(6\text{ cm}) - PA(\text{opt.})\} = -7.2 \pm 2.8^\circ$, and $\Delta\chi$ (2 cm-opt.) = $-0.7 \pm 2.5^\circ$. Interpreting $\Delta\chi$ in terms of the Faraday rotation would result in a rotation measure (RM) of the hotspot $RM = (-35 \pm 13)$ rad m⁻², roughly consistent with the overall rotation measure of 3C33 south: $RM = -12$ rad m⁻² (refs 13, 14). In addition, there seems to be depolarization at 6 cm: $P(6\text{ cm})/P(\text{opt.}) = 0.75 \pm 0.07$ which is not present in the 2-cm data: $P(2\text{ cm})/P(\text{opt.}) = 1.02 \pm 0.07$ (see ref. 6). This cannot be attributed to internal Faraday rotation alone, because even for a homogenous source (that is constant magnetic field and elec-

Table 1 Optical polarization in 3C33 (south) and comparison with radio data

Position	$\Delta\alpha, \Delta\delta$ (arc s)	Optical (660 nm)			Radio (6 cm)		
		Relative flux/beam†	<i>P</i> (%)	<i>PA</i> (deg)	Relative (flux/beam*)	<i>P</i> (%)	<i>PA</i> (deg)
1*	-0.3, -0.3	0.73	34 ± 6	57 ± 7	0.61	29	45
2*	0, 0	1.00	29 ± 4	53 ± 6	1.00	18	50
3*	0.3, 0.3	0.85	23 ± 4	50 ± 7	0.73	?	?
4*	0.4, -0.7	0.50	28 ± 6	44 ± 12	0.42	25	45
5*	-0.7, 0.3	0.72	39 ± 7	63 ± 5	0.24	36	54
6*	-0.7, 0.7	0.45	48 ± 13	70 ± 5	0.12	34	65
7	-1.4, 1.1	0.25	46 ± 25	82 ± 6	0.06	46	75
8	-1.1, 1.4	0.23	48 ± 24	81 ± 7	0.06	46	75
9	-0.7, 1.8	0.21	43 ± 19	78 ± 10	0.06	50	76
10	1.0, 1.1	0.22	27 ± 11	128 ± 23	0.30	20	120
11	1.6, 2.0	0.21	26 ± 11	88 ± 9	0.24	20	90
A-1	-1.5, 1.8	0.56	10 ± 8	91 ± 15	—	—	—
A-2	-2.2, 2.5	2.32	4.8 ± 1.6	78 ± 8	—	—	—

* Optical polarization exceeds 3σ .

† The optical data are integrated within an effective beamwidth of 2.0 arc s FWHM, the beamwidth of the radio data⁶ is 0.85 arc s FWHM.

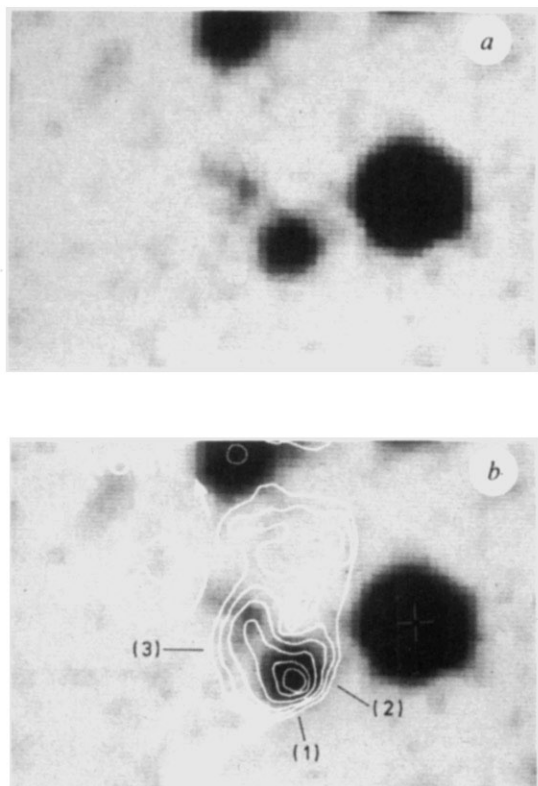


Fig. 2 High signal-to-noise image of 3C33 south at $\lambda = 660$ nm. *a*, Superposition of $3 \times 3 \times 2$ images from the three sets of CCD frames. The pixel size is 0.35×0.35 arc s; 1 arc s corresponds to a linear size of 1.05 kpc. (Secondary images of star A and the hotspot are present near the upper left edge of the field). *b*, As *a* with 5-GHz contours from ref. 6 superimposed. Individual regions are: 1, hotspot; 2, north-west wing; 3, backward tail. The direction towards the central galaxy is indicated by an arrow.

tron density along the line-of-sight) the above RM would yield a depolarization $< 1\%$.

Discussion and interpretation

The detection of a highly polarized optical counterpart for the hotspot in the southern lobe of 3C33 establishes a synchrotron origin for the radiation emitted at optical frequencies. Furthermore, the excellent agreement in both amount and orientation between optical and radio polarization strongly argues for a common emission region. Thus, for the first time, we can base the discussion of the overall nonthermal hotspot spectrum in a classical double radio source on the frequency range 10^9 – 10^{14} Hz.

Calibrating our data with reference to several standard stars observed with the same instrument yields a hotspot flux

$$F_\nu (\nu = 4.5 \times 10^{14} \text{ Hz}) = 4.2 \pm 0.8 \mu\text{Jy}$$

within an aperture of 2 arc s diameter. Based on the mean surface brightness of the diffuse patch as given by Simkin⁷ a blue magnitude $B' = 25.0 \pm 0.5$ would result using the same aperture. Actually this value should be regarded as an upper limit because hardly any blue radiation is detected from the hotspot itself on Arp's 200-in plate (Fig. 1c in ref. 7) and the synchrotron origin of that radiation is not established. Therefore, we assume $F_\nu (\nu = 6.8 \times 10^{14} \text{ Hz}) \leq 0.5 \mu\text{Jy}$ for the blue hotspot flux. A more recent flux measurement based on Arp's 200-in plate supports this value (S. M. Simkin, personal communication).

In Fig. 4, the optical flux values are compared with simple power law extrapolations $F_\nu \sim \nu^{-0.89 \pm 0.05}$ of the radio flux at 5 GHz (ref. 6). Obviously, a smooth link to the optical data requires a sharp spectral break around 10^{14} Hz.

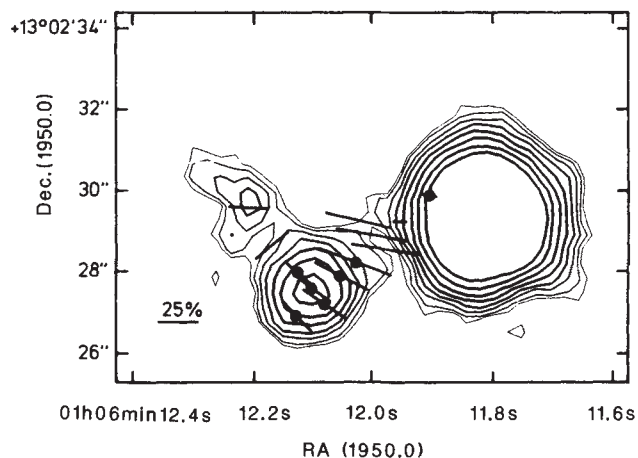


Fig. 3 Map of optical brightness and polarization within 3C33 south. The contours give 10%, 14%, 20%, ..., 80% of the peak flux at 660 nm (1.5 arc s resolution). The vectors superimposed show the fractional amount of linear polarization, integrated over an effective circular beam of 2.0 arc s FWHM. Positions where the amount of polarization exceeds 3σ are indicated by a central dot (see Table 1).

In fact, the synchrotron cutoff spectrum expected from a powerlaw electron energy distribution with a high-energy cutoff at $E_{\text{max}} = \gamma_{\text{max}} m_e c^2$ and a corresponding cutoff frequency¹⁵

$$\nu_c = 46 \text{ Hz } \gamma_{\text{max}}^2 \frac{B_\perp}{1 \text{ nT}} = (1.3 \pm 0.1) \times 10^{14} \text{ Hz}$$

describes the data sufficiently well ($B_\perp$ = magnetic field perpendicular to the line-of-sight). A similar spectrum is observed from the hotspot at the outer end of 3C273A (ref. 11). Infrared photometry at $2.2 \mu\text{m}$ (expected flux: 50 up to $130 \mu\text{Jy}$) of the hotspot in 3C33 south would provide a crucial test for the cutoff spectrum which we will assume throughout the following discussion. So the power emitted by the hotspot between millimetre and infrared wavelengths is double that from the radio domain:

$$L_{\text{mm-IR}} = 3.5 \times 10^{35} \text{ W (frequency range: } 10^{11}\text{--}10^{14} \text{ Hz)}$$

$$L_{\text{radio}} = 1.6 \times 10^{35} \text{ W (frequency range: } 10^8\text{--}10^{11} \text{ Hz)}$$

Additional physical parameters estimated from the canonical formulae^{3,9} for the hotspot and its adjacent regions (see Fig. 2) are summarized in Table 2. Note that only the total luminosity but not the minimum energy density and the corresponding magnetic field B_{me} ($\equiv$ equipartition magnetic field) is affected by the detection of optical synchrotron radiation. Most strongly, however, the required maximum Lorentz factors γ_{max} of the relativistic electrons are increased. Within the hotspot, γ_{max} should be at least 2×10^5 (that is, $E_{\text{max}} = 100$ GeV) or even higher if the magnetic field strength is overestimated by its equipartition value. Only lower limits to γ_{max} can be set in the backward tail and the north-west wing where our optical data are less stringent as regards the high-frequency cutoff. These high Lorentz factors yield very short synchrotron lifetimes of the order of 100 yr (see Table 2). Even if $B_\perp = 0.1 B_{\text{me}}$ the synchrotron lifetimes are too short for electrons accelerated in the hotspot to reach the backward tail or the north-west wing without losing about half of their energy. The cutoff frequency would then be shifted to $\nu_c \leq 4 \times 10^{13}$ Hz, completely depressing the optical radiation. The observed optical brightness distribution, however, does not differ significantly from that found at 5 GHz (see Table 1).

On the other hand, much weaker fields (that is $B_\perp < 0.01 B_{\text{me}}$) are highly implausible, because an overwhelming fraction of the total energy would then be stored in relativistic particles. In this case, the highly ordered magnetic field, required for the

observed radio-to-optical polarization and for fast electron drift velocities $v_{\text{drift}} \approx c$, could not be maintained against destructive induction effects due to the tightly gyrating electrons (or positrons). Therefore, highly relativistic particles must be accelerated inside all the regions from where optical synchrotron radiation is observed. Furthermore, from the similarity between the optical and the 6-cm map ($\tau_{\text{synch}}(6 \text{ cm}) \approx 140\tau_{\text{synch}}(\text{opt.})$) we conclude that the drift velocities of the relativistic electrons through the hotspot region have to be much smaller than the speed of light.

Our observation of spatially resolved optical synchrotron emission demands an acceleration process which can boost electrons to highly relativistic energies and works efficiently over a region of $>2 \text{ kpc}$ in extent. Significant improvement of the angular resolution at optical frequencies (that is, Space Telescope resolution) should show whether this accelerator consists of several well separated substructures or smoothly fills the whole emission region.

With regard to the nature of the acceleration mechanism, several authors^{16,17} have shown that diffuse shock acceleration (Fermi acceleration) at strong shocks would produce power law energy distributions of relativistic electrons with a high-energy cutoff if synchrotron losses are taken into account. Thus, Fermi acceleration at strong shocks, which may be very efficient due to 'bootstrapping' by nonlinear effects¹⁸, seems to be a promising way of explaining the synchrotron spectrum observed from the hotspots in 3C33 south and 3C273A. Note, however, that such a high-energy cutoff is not a unique feature of diffuse shock acceleration but should occur in any process where the energy gain per duty cycle may eventually be balanced by the synchrotron losses (in practice this also limits the maximum energy in the case of man-made electron synchrotrons).

The presence of a strong shock within the hotspot fits well into the standard twin jet model for double radio sources^{19,20} where hotspots represent the working surfaces of collimated supersonic flows ramming into the intergalactic medium. Detailed hydrodynamic simulations of jets and their beam heads were presented by Smith *et al.*²¹. On first sight, there are obvious similarities between the observed morphology of the hotspot region in 3C33 south and the numerically-simulated structures of beam caps (compare our Figs 2 and 3 with Figs 4–8 in ref. 21). Within the framework of the gas dynamical jet models, the high linear polarization²² and the sharp leading edge of the hotspot and the north-west wing (which has a less pronounced south-east correspondence) as well as the detection of a well separated precursor on Dreher's highly resolved 15-GHz map⁵ would argue for inclination angles $30^\circ < \theta_j < 70^\circ$ between the jet

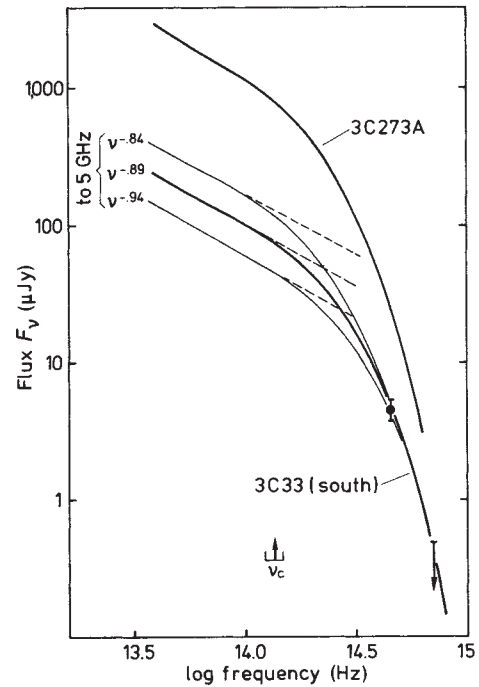


Fig. 4 Comparison between optical flux from the hotspot in 3C33 south (see text) and power law extrapolations $F_\nu \sim \nu^\alpha$ ($\alpha = -0.89 \pm 0.05$) of the flux density at 5 GHz (ref. 6). A synchrotron cutoff spectrum with cutoff frequency ν_c allows a smooth linking between radio spectrum and optical data. For comparison, the spectrum of 3C273 A (hotspot at the leading edge of the jet) is plotted, as it would be observed with the same redshift as 3C33 ($z = 0.06$).

axis and the line-of-sight. Such direct comparison between model jets and observations, however, may be greatly misleading, since Smith *et al.*²¹ did not take into account dissipative losses and magnetic fields, which do have an important role in producing the observed radiation. In addition, their brightness maps²¹ were calculated under the assumption that the path lengths of relativistic particles significantly exceed the dimensions of the beam caps. As discussed above, this assumption is not valid in the case of 3C33 south.

Nevertheless, gas dynamic models of jets and hotspots may have some physical significance, if the gas pressure dominates the total pressure within the jet. Then the jet speed v_j can be

Table 2 Physical parameters of the synchrotron-emitting regions

Region	Hotspot	NW wing	Backward tail
Positions (see Table 1)	1, 2, 3, 4, 5	7, 8	10, 11
Assumed: Volume (kpc^3)	1.0	2.0	1.6
Spectrum $F_\nu \sim \nu^\alpha$, $\alpha =$	-0.89	-0.8	-0.8
Calculated:			
Luminosity L_{synch} (10^{35} W)	5.5	2.4	4.1
Minimum energy density u_{me} (10^{-8} J m^{-3})	3.8	1.1	1.6
Magnetic field B_{me} (nT)	64	34	42
Lorentz factor			
$\gamma_{\text{max}} (B_{\text{me}}, 10^{14} \text{ Hz})$ (10^5)	1.9	$>2.6^\dagger$	$>2.4^\dagger$
$\gamma_{\text{max}} (0.1 B_{\text{me}}, 10^{14} \text{ Hz})$ (10^5)	6.1	$>8.4^\dagger$	$>7.5^\dagger$
Synchrotron lifetime			
$\tau_{\text{synch}} (B_{\text{me}}, 10^{14} \text{ Hz})$ (10^3 yr)	0.09	<0.23	<0.17
$\tau_{\text{synch}} (0.1 B_{\text{me}}, 10^{14} \text{ Hz})$ (10^3 yr)	2.8	<7.4	<5.4
Distance from hotspot d (kpc)	—	1.5	2.5
$d/c\tau_{\text{synch}} (B_{\text{me}}, 10^{14} \text{ Hz})$	$18^\ddagger$	>21	>48
$d/c\tau_{\text{synch}} (B_{\text{me}}, 5 \text{ GHz})$	$0.12^\ddagger$	>0.14	>0.34

* We assume $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = \phi$, $\nu_{\text{min}} = 10^7 \text{ Hz}$, $\nu_{\text{max}} = 10^{14} \text{ Hz}$, relative amount of energy stored in protons: $k = 100$, see ref. 9.

† No evidence for a cutoff at $\nu_c = 10^{14} \text{ Hz}$.

‡ Calculated for $d = 0.5 \text{ kpc}$ (that is, HWHM of hotspot profile).

estimated from the requirement that the bulk kinetic energy supplied by the jet flow should at least balance the emitted power of the lobe. If we estimate the jet density from the rotation measure of the hotspot by assuming internal Faraday rotation in a magnetic field $B_{\parallel} = B_{\text{me}}/\sqrt{3}$, we find $v_j \geq 0.14 c$ (for a cold jet with radius $R_j = 0.5$ kpc, $L \geq 10^{36}$ W and $\epsilon = 0.1$ see ref. 23). An alternative argument based on the momentum flux as calculated from the minimum pressure $p_{\text{me}} = u_{\text{me}}/3$ (Table 2) yields a similar lower limit for the jet speed²⁴. So the power emitted from the southern hotspot and lobe of 3C33 could be supplied by a jet flow with mildly relativistic speeds as are directly observed in the galactic twin jet source SS433 (ref. 25). For comparison, we note that the momentum flux argument yields $v_j > 0.76 c$ in the case of the more powerful double radio source Cygnus A (refs 24, 26).

CCD polarimetry, therefore, opens up a new frequency range for investigating the nonthermal radiation from radio lobes. Our results from 3C33 south point to several general conclusions.

First, optical synchrotron radiation from hotspots in extended radio lobes inevitably demands the presence of *in situ* acceleration processes. In 3C33 south, the synchrotron-emitting region is clearly extended, thus the acceleration process must occur within a large volume ($d \geq 2$ kpc). Space Telescope observations should allow us to decide whether the acceleration takes place

within more compact substructures.

Second, our data on the hotspots in 3C33 south and 3C273A indicate synchrotron spectra with cutoffs at nearly identical frequencies $\nu_c = 10^{14}$ Hz. We note that very similar spectra are observed in some QSO cores^{27,28}. Although Fermi acceleration at strong shocks could account for the cutoff, the intriguing fact of similar cutoff frequencies in compact cores ($d < 1$ pc) and extended sources ($d > 1$ kpc) still needs to be explained. Both hotspots emit a substantial fraction of their luminosity in the millimetre to infrared range. This makes observations with IR-arrays very promising.

Finally, any detailed comparison between observations and numerical simulations of jets and hotspots has to be treated with caution, as the processes which are responsible for the observed radiation are not considered in the model calculation (for example, dissipative losses, magnetic fields, and particle acceleration). It is necessary to investigate how these processes affect the hydrodynamic models. As we have found that optical (and possibly radio) maps mainly represent the accelerating regions, detailed models of Fermi acceleration or alternative mechanisms are urgently needed.

This work was based on observations made at the European Southern Observatory, La Silla (Chile) with the 2.2-m telescope of the Max Planck Society.

Received 30 July; accepted 24 October 1985.

- Hargrave, P. J. & McEllin, M. *Mon. Not. R. astr. Soc.* **173**, 37–63 (1975).
- Simkin, S. M. *Astrophys. J.* **234**, 56–67 (1979).
- Miley, G. K. A. *Rev. Astr. Astrophys.* **18**, 165–218 (1980).
- Gopal-Krishna, Joshi, M. N. & Ananthakrishnan, S. *Astrophys. Lett.* **17**, 11–14 (1976).
- Dreher, J. W. *Astr. J.* **86**, 833–847 (1981).
- Rudnick, L., Saslaw, W. C., Crane, P. & Tyson, J. A. *Astrophys. J.* **246**, 647–652 (1981).
- Simkin, S. M. *Astrophys. J. Lett.* **222**, L55–L58 (1978).
- Crane, P., Tyson, J. A. & Saslaw, W. C. *Astrophys. J.* **265**, 681–700 (1983).
- Pacholczyk, A. G. *Radio Astrophysics* (Freeman, San Francisco, 1970).
- Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. Mod. Phys.* **56**, 255–351 (1984).
- Röser, H.-J. & Meisenheimer, K. *Astr. Astrophys.* (in the press).
- Newell, E. B. in *Image Processing in Astronomy* (eds Sedmak, G., Capaccioli, M., Allen, R. J.) (Osservatorio Astronomico di Trieste, 1979).
- Haves, P. *Mon. Not. R. astr. Soc.* **173**, 553–568 (1975).
- Vallée, P. J. & Kronberg, P. P. *Astrophys. J.* **193**, 303–308 (1974).

- Ginzburg, V. L. & Syrovetskij, S. I. *The Origin of Cosmic Rays* (transl. Massey, H. S. H.) (Pergamon, Oxford 1964).
- Webb, G. M., Drury, L. O'C. & Biermann, P. *Astr. Astrophys.* **137**, 185–201 (1984).
- Bregman, J. N. Preprint (1984).
- Völk, H. J. in *Trends in Physics Vol. 1*, 162–169 (European Physics Society 1985).
- Rees, M. J. *Nature* **229**, 312–317 (errata p. 510) (1971).
- Blandford, R. D. & Rees, M. J. *Mon. Not. R. astr. Soc.* **169**, 395–415 (1974).
- Smith, M. D., Norman, M. L., Winkler, K.-H. A. & Smarr, L. *Mon. Not. R. astr. Soc.* **214**, 67–85 (1985).
- Laing, R. A. *Mon. Not. R. astr. Soc.* **193**, 439–449 (1980).
- Perley, R. A., Bridle, A. H. & Willis, A. G. *Astrophys. J. Suppl.* **54**, 291–334 (1984).
- Dreher, J. W. in *Physics of Energy Transport on Extragalactic Radiosources* (eds Bridle, A. H. & Eilek, J. A.) 109–113 (NRAO Proc., Green Bank, 1984).
- Margon, B. A. *Rev. Astr. Astrophys.* **22**, 507–536 (1984).
- Perley, R. A., Dreher, J. W. & Cowan, J. J. *Astrophys. J. Lett.* **285**, L35–L38 (1984).
- Bregman, J. N. *et al. Nature* **293**, 714–717 (1981).
- Rieke, G. H., Lebofsky, M. J. & Wisniewsky, W. Z. *Astrophys. J.* **263**, 73–78 (1982).

Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins

Frank D. McKeon*, Marc W. Kirschner* & Daniel Caput†‡

* Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, California 94143, USA

† Chiron Corporation, 4560 Horton St, Emeryville, California 94608, USA

The A, B and C lamins are the major proteins of the nuclear envelope. The complete nucleotide sequence of the coding region of the A and C lamins shows that these proteins are identical except for their carboxy termini. The most prominent structural feature of both lamins is an α -helical region of repeating heptads of amino acids that shows striking homology with the entire family of cytoplasmic intermediate filament proteins. These features suggest that the nuclear envelope is made up of a network of coiled-coil polymers.

THE eukaryotic nuclear envelope is responsible for segregating the cytoplasm from the nucleoplasm and thereby channelling all transport across the nuclear membrane through an array of specialized structures termed nuclear pores^{1,2}. As the nuclear envelope is the site at which the very first events of chromosome condensation occur^{3,4}, it has also been suggested that it is involved in the maintenance of interphase chromosome organization⁵. Despite the fact that the nuclear envelope appears to be

very stable in isolated nuclei, it can be very dynamic when it breaks down and rapidly reforms during mitosis^{6,7}.

Although the nuclear envelope is not completely defined biochemically, it is now known that its major structural component is the nuclear lamina, a durable, fibrous layer on the nucleoplasmic surface of nuclear membranes⁸ which acts as a scaffold for the nuclear pores and may interact with interphase chromatin^{9,10}. The structural integrity of the lamina is under strict cell-cycle control, as shown by its disintegration soon after chromosome condensation in prophase and its reappearance around telophase chromosomes^{6,7}.

‡ Present address: ELF Biorecherche, La bege-innople, BP 137 31328, Castanet Tolosan Cedex, France.

In mammals the nuclear lamina is composed primarily of three proteins, lamins A, B and C, present in equal amounts in the interphase nuclear envelope¹¹. Of these proteins, lamins A (relative molecular mass, M_r = 70,000) and C (M_r = 60,000) appear to be closely related on the basis of peptide maps and immunological cross-reactivity^{11,12}. By these criteria, lamin B (M_r = 68,000) has a very different structure. Increased phosphorylation of the lamin proteins occurs before nuclear envelope disintegration in prophase and probably plays a crucial role in the regulation of associations between lamin proteins^{11,13,14}.

To understand better the structure of the nuclear envelope, we have examined in detail the structure of the lamin proteins. We wish to know whether these proteins interact directly with the chromatin or with the nuclear membrane, and also what accounts for the considerable stability of the nuclear envelope in interphase and yet allows for its complete disassembly and reassembly during mitosis. In addition there are questions of the individual role of the A, B and C lamins, their organization and their mechanism of assembly. As a first step in this analysis, we have isolated complementary DNAs containing the complete coding sequences of both the A and C lamins, screened using a high-titre anti-lamin serum obtained from a patient with the autoimmune disease linear scleroderma¹⁵. The predicted amino-acid sequences of the A and C lamins are indicative of unusual secondary structures and show a strong homology to all five classes of cytoplasmic intermediate filament proteins¹⁶⁻¹⁸. These structural features suggest that the nuclear envelope derives its structural integrity from coiled-coil interactions between lamin proteins in a manner similar to those in the intermediate filaments.

cDNAs encoding the A and C lamins

Human autoantibodies to the A and C lamins were used to screen a T-cell cDNA expression library. The auto-antiserum (LS-1) was derived from a patient with linear scleroderma¹⁵ and shown to possess high-titre antibodies against conserved determinants of the A and C lamins. The cDNA library was constructed from human T-cell messenger RNA using the primer-adaptor method¹⁹, which ensures that the cDNA is in the proper orientation for expression from the *lacZ* promoter on pUC9 (ref. 20) (Fig. 1a). Bacterial colonies expressing the *lacZ*-lamin fusion proteins were detected using the LS-1 serum, followed by ¹²⁵I-goat anti-human immunoglobulin G antibodies²¹. Although we expected only 1 of 3 cDNAs to be in the proper reading frame for expression, we found that 14 out of 15 of the A and C lamin cDNAs in the library could be detected using the antibodies. We suspect that the G-C tail between the *lacZ* gene and the cDNA coding sequence caused some frameshifting, resulting in this unusually high proportion of cDNAs expressing immunoreactive proteins.

The 15 putative lamin cDNAs obtained from this library fell into two groups, based on restriction mapping, distinguished by different 3' ends. The longest cDNA from each group was expressed in *Escherichia coli* and the fusion proteins detected by Western blots²² (Fig. 1b). The cDNA-7 expresses a protein with an apparent M_r of 72,000 whereas cDNA-13 yields a fusion protein of M_r ~84,000. As shown below, these two cDNAs show extensive homologies over most of their sequences. Although these apparent M_r s are considerably larger than the native lamin standards (Fig. 1b), we explain this by the presence of a 5'-untranslated sequence in-frame with the *lacZ* initiating methionine. To determine whether these immunoreactive proteins are indeed lamin proteins, we examined proteolytic digests of the fusion proteins produced in *E. coli* from the cDNAs together with the native lamin proteins derived from nuclear envelopes. The native lamin C and fusion protein from a putative lamin cDNA yield very similar peptide patterns after partial digestion with V8 protease (Fig. 2a). We obtained further evidence that these cDNAs encode nuclear lamins by observing the immune response in mice to the injected fusion proteins. Four weeks

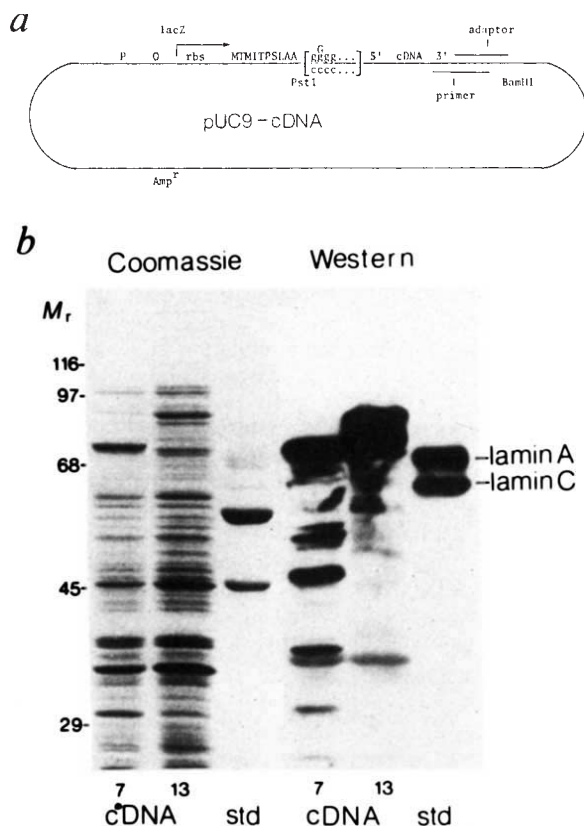


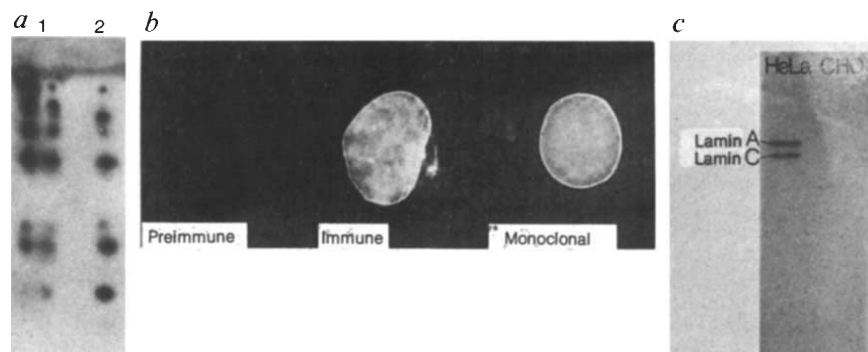
Fig. 1 Construction and expression of cDNAs. *a*, cDNA was synthesized from poly(A) RNA obtained from peripheral human lymphocytes using a primer-adaptor method¹⁹. This method constructs the recombinants such that the cDNA is in the proper orientation with respect to the *lacZ* start of transcription. The resulting fusion protein starts with the 10 amino acids of the β -galactosidase protein (MTMITPSLAA), continues with ~5 glycine residues translated from the G-C tail, and ends with the cDNA sequence. P, promoter; O, operator; rbs, ribosome binding sequence; adaptor and primer are both synthetic oligonucleotides as described elsewhere¹⁹. *b*, Expression of cDNA proteins in *E. coli*. The figure shows a Coomassie blue-stained SDS-polyacrylamide gel⁴³ showing solubilized whole *E. coli* (D1210) harbouring cDNAs 7 and 13 after 10 h of induction with 2 mM isopropyl- β -D-thiogalactopyranoside (IPTG). A Western blot²² identical to the stained gel was probed with the human autoantibody to the A and C lamins followed by ¹²⁵I-goat anti-human IgG. The fusion proteins have a higher M_r than those of the lamin standards derived from CHO interphase nuclear envelopes.

after immunization with the cDNA-7 fusion protein, we detected a strong immune response to the nuclear envelope in the serum using an immunofluorescence assay (Fig. 2b). Spleen cells from one mouse were fused with myeloma cells to produce monoclonal antibody-secreting hybridomas²³, of which several recognized the nuclear envelope in the same immunofluorescence assay (Fig. 2b). One of these hybridomas produced antibodies that reacted with A and C lamins in primates only, as shown in the Western blot against both Chinese hamster ovary cell and HeLa cell nuclear envelopes (Fig. 2c). The similarity of the peptide patterns and the highly specific immune response against primate lamins suggest that cDNA-7 and -13 encode the human lamin proteins.

Relationship between A and C lamins

Evidence that cDNA-7 encodes lamin C and cDNA-13 lamin A is provided by Northern blots and DNA sequencing. Laliberte *et al.*²⁴ showed that a 3-kilobase (kb) mRNA encoded a precursor to lamin A of M_r 74,000, whereas a 2.2-kb species encoded lamin C. In agreement with these data, we found that DNA

Fig. 2 Peptide mapping and immunological response to fusion proteins. **a**, Peptide comparison between the cDNA-4 fusion protein and the CHO C lamins. *E. coli* harbouring cDNA-4 and nuclear envelopes from Chinese hamster ovary cells were fractionated on an SDS-polyacrylamide gel⁴³ (cDNA-4 encodes ~55K of lamin C and is identical in sequence to 1,800 nucleotides of cDNA-7. It is expressed at very high levels in *E. coli*, facilitating its use in peptide mapping.) Lamin and fusion proteins were excised and digested with V8 protease⁴⁴. The peptides were then fractionated on an SDS-polyacrylamide gel and transferred electrophoretically to nitrocellulose²². The blot was probed with the LS-1 anti-lamin serum and then with ¹²⁵I-goat-anti-human IgG. Lane 1, lamin C peptides; lane 2, peptides from the cDNA-4 fusion protein. **b**, Immunological response to lamin fusion proteins. D1210 *E. coli* (5 ml) harbouring cDNA-4 was grown overnight in Luria Broth media with 100 µg ml⁻¹ carbenicillin and 2 mM IPTG, pelleted by centrifugation and lysed with 400 µl 6 M guanidine-HCl for 1 min. The lysate was diluted with 1.6 ml H₂O and precipitated by addition of 8 ml acetone. After centrifugation, the pellet was sonicated in an equal volume of Freund's incomplete adjuvant and 20 µl of the emulsion was injected into each footpad of a BALB/c female mouse on day 1, followed by 50-µl injections (interperitoneal) on days 12 and 21. The mice were bled on day 28 and the serum diluted 1:500 for indirect immunofluorescence against BSC1 (African green monkey kidney) cells grown on coverslips and fixed with 1% formaldehyde for 5 min¹⁵. The preimmune serum shows low diffuse reactivity, whereas the immune serum yields a strong response to the nuclear envelope. Analysis of the hybridomas produced from the spleen cells of these mice revealed several that secreted antibodies to the nuclear envelope, at least one of which recognized primate A and C lamins exclusively, as shown by immunofluorescence and Western blotting. **c**, Western blot of HeLa and CHO nuclear lamins probed with a monoclonal antibody to the lamin fusion proteins. Equal amounts of HeLa and CHO nuclear lamins were prepared as described elsewhere¹⁵, fractionated on an SDS-polyacrylamide gel and transferred electrophoretically to nitrocellulose.



probes made from either cDNA-7 or cDNA-13 detect a 3- and 2.2-kb mRNA from human cells (Fig. 3). From the respective sizes of the cDNAs, cDNA-13 (2.7 kb) must encode lamin A and cDNA-7 (2.2 kb) lamin C; in fact, cDNA-13 encodes a polypeptide ~11,000 (11K) larger than the cDNA-7 peptide produced in *E. coli* (Fig. 1b). The sequences of the cDNA-7 and the cDNA-13 clones (Fig. 4) show that they are identical from the extreme 5' end up to a position near the 3' end. When translated, the pre-lamin A sequence of cDNA-13 encodes a protein 12.5K larger than the cDNA-7 sequence, in good agreement with the products observed in gels.

The exact sizes of the pre-A and C fusion products in *E. coli* (88,021 and 73,757, respectively, determined from the sequence) are larger than those expected for the pre-A and C lamin proteins. The reason for this was apparent from the sequence. Both cDNAs have 5' noncoding regions in-phase with the initiating methionine of β -galactosidase. Therefore, ~9K of amino-acid sequence read from the noncoding 5' end of the lamin mRNAs is present in the fusion products. In examining the sequences it was clear that the pre-A and C lamins must be initiated from a single in-frame methionine at the position shown in Fig. 4, which results in a M_r of 79,409 for the pre-A lamin and 65,145 for the C lamin. The pre-A lamin is further processed in the nucleus to give the mature A lamin^{14,24}. It is not possible to determine the site of proteolytic cleavage from the nucleic acid sequence alone.

Homology to intermediate filaments

Over a large part of the amino-acid sequence, lamins A and C show strong homology with intermediate filament proteins. There are 111 identities out of 394 possible matches with only 2 gaps when lamins A and C are compared with vimentin using a sequence-matching computer program²⁵. The sequence and the structural homologies between lamins and intermediate filaments are illustrated using the suggested secondary structure of lamin C (Fig. 5).

Intermediate filaments contain a ~310-residue stretch of amino acids which is thought to be α -helical and which has a distinctive structural feature suggesting that these helices can form coiled-coils with a similar α -helix¹⁶⁻¹⁸. This distinctive feature is the presence of repeats of seven amino acids, such that if the positions in these repeats are denoted a, b, c, d, e, f

and g, residues a and d are generally hydrophobic and this sequence is repeated without gaps. This places the a and d residues next to each other on one side of the α -helix so that two helices can wind around each other, held together by hydrophobic interactions. Tropomyosin is the best-studied example of a protein with extensive helical domains capable of forming coiled-coils²⁶. Intermediate filaments contain three or four major α -helical regions spaced by linkers which do not form these heptad repeats¹⁶⁻¹⁸.

As shown in Fig. 5, a short stretch of amino acids, which has a high tendency to form an α -helix but does not have heptad repeats¹⁸, precedes the first coil. Directly after this sequence is the coil 1A domain of 39 residues, in which all 11 a and d positions in the heptad repeats are hydrophobic. The heptad repeats are shown directly in Fig. 4 where the a and d positions are placed in the corresponding position used for intermediate filament proteins¹⁶. The coil 1A region and the preceding helix are not only structurally homologous but also closely related in sequence. In Fig. 5b the precoil helix and amino-terminus of

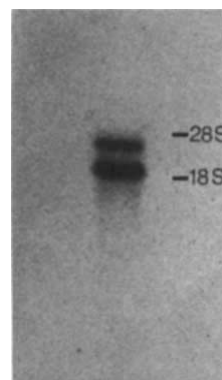
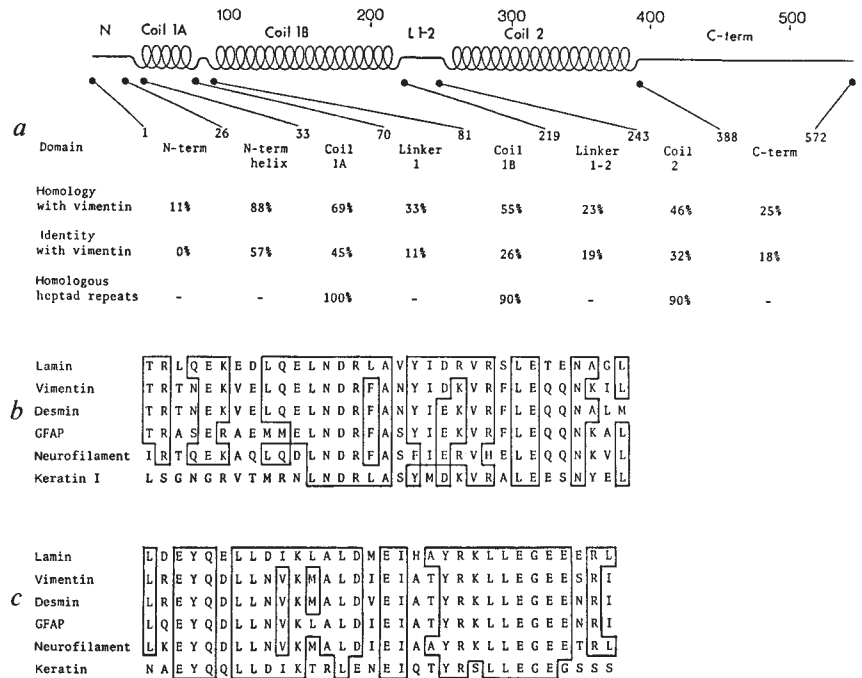


Fig. 3 The mRNA complexity of A and C lamins. Human poly(A) selected RNA was fractionated by electrophoresis through a 1% agarose gel containing formaldehyde and transferred to nitrocellulose⁴⁵. This filter was probed with a nick-translated 1-kb *Hind*III-*Bam*HI fragment from the 3' end of cDNA-7, yielding two bands migrating at 2.2 and 3.0 kb. The 28S and 18S rRNAs are 5.0 and 2.1 kb.

Fig. 5 *a*, Secondary structure and sequence homologies of lamins and intermediate filaments. The lengths of various coil, linker and non-helical domains are drawn to scale and aligned with intermediate filament structures¹⁶. The homology arises from using structurally related amino acids of the following families: A, I, L, V; Q, N; F, Y, W; P; E, D; H, R, K; S, T; G; C; M. Identities were from alignment matching²⁵. Homologous heptad repeats refer to hydrophobic residues or polar substitutions conserved in vimentin and lamin C in the *a* and *d* positions of the heptads. *b*, Sequence comparison of the amino-terminal region of coil 1A of lamins A and C (residues 27–58) with those of hamster vimentin³⁷, chicken desmin³⁹, mouse GFAP, porcine neurofilament³⁸ and mouse epidermal keratin II (ref. 40). *c*, Sequence comparison of lamins A and C with the carboxy-terminal region of coil 2 (residues 355–387).



repeats and the α -helical structure of lamins A and C end abruptly at residue 387. In coils 1A, 1B and 2, the *a* and *d* positions contain 85% uncharged residues, whereas 48% of the remaining positions are uncharged. The corresponding values for mouse epidermal keratin are 81 and 61% (ref. 27).

We predict that the amino- and carboxy-terminal domains have mixed secondary structure. These domains show no overall homology with intermediate filaments. The very short amino-terminal non-helical region is strongly basic in lamins and in vimentin. The much larger carboxy-terminal region has clustered charged groups; in particular, 24% of the 50 amino acids next to coil 2 are basic, with an unusual stretch containing 30% serine residues. The extreme carboxy-terminus of the C lamins is highly charged.

To give some statistical picture of where the A and C lamins fit in the intermediate filament family, we have compared the sequences using a sequence-matching algorithm²⁵. There is a very strong relationship of lamins to all intermediate filaments, except for the 68K neurofilament protein (Table 1). These relationships are weaker than the close relationship of vimentin, desmin and glial fibrillary acidic protein (GFAP) (three inter-

mediate filament proteins considered to be members of a related superclass¹⁷) to each other. However, the relationship of the A and C lamins to the vimentin/desmin/GFAP group is much closer than the relationship of this superclass to the neurofilament proteins. To show that these relationships are not merely a consequence of the α -helical structure involved in coiled-coil interactions, we have compared all of these sequences with rabbit skeletal muscle tropomyosin and nematode myosin. The structural identity here, although possibly statistically significant, is much lower than the lamin-intermediate filament comparisons (Table 1).

Implications of lamin structure

Consideration of the A and C lamins as members of the intermediate filament protein family depends on their possession of several characteristic features of that family. The most striking structural features are: the long α -helical stretch with heptad repeats; the conserved position of the linkers; the conserved replacement of charged amino acids in the heptad repeats; the conserved position of the phase changes in coil 2; and, particularly, the strong regions of sequence identity in coils 1A and

Table 1 Alignment identities between amino-acid sequences of proteins capable of coiled-coil interactions

	Lamin C	Vimentin	GFAP	Desmin	Keratin II	Keratin I	Neurofilament (68K)	Myosin heavy chain
Vimentin	44	—	—	—	—	—	—	—
GFAP	43	201	—	—	—	—	—	—
Desmin	41	123	159	—	—	—	—	—
Keratin II	38	62	68	69	—	—	—	—
Keratin I	34	62	70	68	55	—	—	—
Neurofilament	18	29	33	36	33	33	—	—
Myosin heavy chain	9	4	6	6	4	12	3	—
Tropomyosin	5	5	2	5	2	8	3	17

Sequences were aligned using an alignment score program²⁵ based on the Needleman and Wunsch algorithm³⁶. The alignment score is the number of standard deviations (s.d.) by which the best alignment score differs from the mean alignment score of the randomly permuted sequences, divided by the s.d. of the distribution. The sequence of lamin C is compared with that for hamster vimentin³⁷, mouse GFAP³⁸, chicken smooth muscle desmin³⁹, human type II epidermal keratin²⁸, mouse type I epidermal keratin²⁷, porcine 58K neurofilament protein⁴⁰, nematode myosin heavy chain⁴¹, and rabbit skeletal muscle tropomyosin⁴².

2. Although there is strong sequence divergence in the carboxy-terminus of lamins A and C, this is also a common feature of intermediate filaments, for example, the 50K and 56K keratins²⁷⁻²⁹. The fact that the A and C lamins share a 1,910-residue stretch of identical nucleotide sequence suggests strongly that these proteins arise from the same gene by differential processing. In support of this suggestion, we see only a single band in *Eco*RI and *Bam*HI digests of human genomic DNA (data not shown). A full understanding of how the mRNA sequences arise requires a study of genomic clones. There are no obvious sequence homologies with other known proteins in the carboxy-terminal and amino-terminal non-coil domains of the lamins. In particular, there is no extensive hydrophobic region characteristic of membrane-binding domains. The membrane anchorage sites may be a property of lamin B as suggested by biochemical studies¹¹.

In analysing lamin structure, we were fortunate that the proteins to which they are homologous have been studied so extensively. Despite considerable sequence divergence outside the central α -helical domain, all the intermediate filament proteins seem to be built on the same model. The basic substructure of intermediate filaments is a coiled-coil molecule thought to be composed of two chains¹⁶⁻¹⁸. Despite the presence of heterogeneous amino- and carboxy-terminal domains, all intermediate filament proteins assemble into the same fundamental 10-nm filament, but the assembly is thought to involve interactions in the non-coiled-coil regions. As non-helical domains in lamins are not related to those in intermediate filament proteins, we cannot be sure that the lamins will assemble into 10-nm filaments. Thus, nuclear lamins almost certainly form coiled-coil oligomers that probably assemble further into elongated molecules of as yet undetermined structure. (Note that Goldman and co-workers have reported that putative lamin-like proteins from BHK cells will form paracrystals with the same axial repeat as found in intermediate filaments³⁰. This and other preliminary information suggested to these workers that lamins and intermediate filaments might share structural homologies³¹.) Unfortunately, a definitive structure of the lamina has not been observed using the electron microscope. We could postulate that lamin filaments are arranged in long fibres or in a dense network similar to that of spectrin in the erythrocyte plasma membrane³². The 1:1:1 stoichiometry of the A, B and C lamins in the nuclear envelope suggests that lamin

B also contains an extensive α -helical region capable of forming a coiled-coil with the lamin A or C. Thus, the nuclear lamina, like intermediate filaments, may comprise a heteropolymer (by analogy with the keratins and neurofilaments, where differences between the A/C and B lamins may be principally in the non-helical regions). As in neurofilaments, there may be some flexibility in the exact composition of these heteropolymers³³.

The recent demonstration that the nuclear envelope of *Xenopus* oocytes and early embryos is composed of only one form of nuclear lamin protein suggests that a homopolymer of lamin molecules is sufficient to provide the stability of the nuclear envelope^{34,35}. The introduction of additional types of lamin proteins, which occurs at a specific developmental step (the mid-blastula transition in *Xenopus*), indicates that more complex functions of the envelope are required in nuclei undergoing active transcription^{34,35}.

It remains a fascinating structural question as to how such a stable and chemically resistant structure involving coiled-coil interactions over a distance of 350 amino acids could be destabilized by phosphorylation during mitosis^{11,13,14}. It will be of interest to determine whether the phosphorylation sites affecting lamin stability are located in the coiled-coil domains or whether they are found in the amino or carboxy termini. Although the α -helical domains may, as in the intermediate filaments, form the basic structure of the nuclear envelope, it may be the non- α -helical domains on the amino and carboxy termini that are involved in supporting the nuclear pores and chromatin, and interacting with other nuclear proteins. An investigation of these functions should provide insights into various aspects of the structure and activity of the nucleus.

We thank Pablo Garcia, Fernando Bazan and Robert Fletterick for help with the sequence homology search and the secondary structure predictions, David Drechsel for the peptide mapping experiments and help in the early stages of this project; Sumire Kobayashi for technical assistance; Jeff Edman for instruction on M13 subcloning; Phil Barr and Guy Mullenbach for providing synthetic oligonucleotides; and Ray Sanchez for help with sequencing. We also thank David Gard and Peter Walter for helpful comments on the text; and Cynthia Hernandez for editing the manuscript. F.D.M. is a Giannini Fellow of the Bank of America. This work was supported by grants from the United Scleroderma Foundation, the NIH and the American Cancer Society.

Received 8 August; accepted 3 December 1985.

1. Franke, W. W., Scheer, U., Krohne, G. & Jarasch, E.-D. *J. Cell Biol.* **91**, 398-50s (1981).
2. Maul, G. G. *Int. Rev. Cytol. Suppl.* **6**, 75-185 (1977).
3. Gurley, L. R., D'Anna, J. A., Barham, S. S., Deaven, L. L. & Tobey, R. A. *Eur. J. Biochem.* **84**, 1-15 (1978).
4. Foe, V. E. & Alberts, B. M. *J. Cell Sci.* **61**, 31-70 (1983).
5. Gerace, L. & Blobel, G. *Cold Spring Harb. Symp. quant. Biol.* **46**, 967-978 (1982).
6. Krohne, G., Franke, W. W., Ely, S., D'Arcy, A. & Jost, E. *Cytobiology* **18**, 22-38 (1978).
7. Gerace, L., Blum, A. & Blobel, G. *J. Cell Biol.* **79**, 546-566 (1978).
8. Fawcett, D. W. *Am. J. Anat.* **119**, 129-146 (1966).
9. Aaronson, R. P. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1007-1011 (1975).
10. Dwyer, N. & Blobel, G. *J. Cell Biol.* **70**, 581-591 (1976).
11. Gerace, L. & Blobel, G. *Cell* **19**, 277-287 (1980).
12. Shelton, K., Higgins, D., Cochran, D., Ruffolo, D. J. & Egle, P. *J. biol. Chem.* **255**, 10976-10983 (1980).
13. Miake-Lye, R. & Kirschner, M. W. *Cell* **41**, 165-175 (1985).
14. Ottoviano, Y. & Gerace, L. *J. biol. Chem.* **260**, 624-632 (1985).
15. McKeon, F. D., Tuffanelli, D. L., Fukuyama, K. & Kirschner, M. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4374-4378 (1983).
16. Weber, K. & Geisler, N. in *Cancer Cells 1, The Transformed Phenotype*, 153-159 (Cold Spring Harbor Laboratory, New York, 1984).
17. Steinert, P. M., Steven, A. C. & Roop, D. R. *Cell* **42**, 411-419 (1985).
18. Traub, P. *Intermediate Filaments* (Springer, Berlin, 1985).
19. Caput, D. *et al. DNA* (in the press).
20. Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
21. Helfman, D. M., Feamisco, J. R., Fiddes, J. C., Thomas, G. P. & Hughes, S. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 31-35 (1983).
22. Burnette, W. N. *Analyt. Biochem.* **112**, 195-203 (1981).
23. Kohler, G. & Milstein, C. *Eur. J. Immun.* **6**, 511-519 (1976).
24. Laliberte, J. F. *et al. J. Cell Biol.* **98**, 980-985 (1984).
25. Orcutt, B. C., Dayhoff, M. O. & Baker, W. C. *Protein Identification Resources* (National Biomedical Program, Washington, D.C., 1984).
26. McLachlan, A. D. & Stewart, M. J. *molec. Biol.* **98**, 293-304 (1975).
27. Steinert, P. M., Rice, R., Roop, D. R., Trus, B. L. & Steven, A. C. *Nature* **302**, 794-800 (1983).
28. Hanukoglu, I. & Fuchs, E. *Cell* **31**, 243-252 (1982).
29. Hanukoglu, I. & Fuchs, E. V. *Cell* **33**, 915-924 (1983).
30. Goldman, R. D. *et al. J. Cell Biol.* **97**, 221a (1983).
31. Zackroff, R. V., Goldman, A. E., Jones, G. R., Steinert, P. M. & Goldman, R. D. *J. Cell Biol.* **98**, 1231-1237 (1984).
32. Bennett, V. A. *Rev. Biochem.* **54**, 273-304 (1985).
33. Shaw, G. & Weber, K. *Expl. Cell Res.* **136**, 119-125 (1981).
34. Benavente, R., Krohne, G. & Franke, W. W. *Cell* **41**, 177-190 (1985).
35. Strick, R. & Hausen, P. *Cell* **41**, 191-200 (1985).
36. Needleman, S. B. & Wunsch, G. D. *J. molec. Biol.* **48**, 443-453 (1970).
37. Quax-Jenken, Y. E. F. M., Quax, W. J. & Bloemendal, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3540-3552 (1983).
38. Lewis, S. A., Balcarek, J. M., Krek, V., Shelanski, M. L. & Cowan, N. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2743-2746 (1984).
39. Geisler, N. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4120-4123 (1981).
40. Geisler, N., Kaufmann, E., Fischer, S., Plessman, V. & Weber, K. *EMBO J.* **2**, 1295-1302 (1983).
41. McLachlan, A. D. & Kairn, J. *Nature* **299**, 226-231 (1982).
42. Stone, D. & Smillie, L. B. *J. biol. Chem.* **253**, 1137-1148 (1978).
43. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
44. Cleveland, D., Fischer, S., Kirschner, M. & Laemmli, U. J. *biol. Chem.* **252**, 1102-1106 (1977).
45. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).
46. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, J. H. *J. molec. Biol.* **143**, 161-178 (1980).

Quasi-periodic oscillations in the X-ray flux of Cyg X-2

G. Hasinger*, A. Langmeier*, M. Sztajno*, J. Trümper*,
W. H. G. Lewin†§ & N. E. White‡

* Max-Planck-Institut für Extraterrestrische Physik,
8046 Garching, FRG

† Massachusetts Institute of Technology, Center for Space Research,
Cambridge, Massachusetts 02139, USA

‡ Exosat Observatory, Space Science Department of ESA, ESOC,
6100 Darmstadt, FRG

We have discovered quasi-periodic X-ray oscillations (QPOs) (30–45 Hz) and low-frequency noise (below 15 Hz) in Cyg X-2 (ref. 1). The centroid frequency of the oscillations is correlated with the source intensity in the same way as in GX 5-1 (refs 2, 3): the frequency increases when the source intensity goes up. Cyg X-2 differs, however, from GX 5-1 (refs 2, 3), where the amplitude of the oscillations and of the low-frequency noise go hand in hand, in that the amplitude of the QPOs decreases while that of the low-frequency noise increases with increasing source intensity. We describe here the spectrum of the persistent emission as the sum of a variable 1.2-keV black-body component, a more or less constant 7 keV thermal bremsstrahlung component and an iron emission line. This is the first report of spectral variations associated with QPOs^{2–6}, which could themselves reflect variations in either the black body or the thermal bremsstrahlung component.

Observations of Cyg X-2 (4U 2142+380) with the Exosat instruments termed the medium energy experiment⁷ (ME) and gas scintillation proportional counter⁸ (GSPC) were carried out during a 14 h period beginning at 02:07 UT on 23 July, 1984. One half of the ME monitored the background. ME data were telemetered with 0.3-s time resolution in 32 energy channels; for 12 of the 14 h high-time-resolution data (7.8 ms) with no spectral resolution were also transmitted. The source intensity varied between ~700 and 1,000 counts per s (c.p.s.) in the ME, corresponding to 1–20 keV luminosities of $(0.9\text{--}1.3) \times 10^{38} \text{ erg s}^{-1}$, respectively for a distance of 8 kpc (ref. 9).

We independently Fourier-analysed 2-s stretches of high-time-resolution data and averaged them to give mean power spectra for four intensity intervals (Table 1). Our normalization¹⁰ was the same as that of Van der Klis *et al.*³. Analysis of shorter or longer stretches of data (0.25–8 s) give the same results. The mean power spectra (Fig. 1) show three components: (1) low-frequency noise (LFN) below 15 Hz; (2) a broad peak characteristic of quasi-periodic oscillations (QPO), whose centroid frequency rises from 30 to 45 Hz as the source intensity increases; and (3) white noise from poissonian counting statistics. The function³:

$$P(f) = 2.0 + A/B \exp(-f/B) + C/\pi E(1 + ((f-D)/E)^2)^{-1}$$

fits the power spectra (solid lines in Fig. 1). The first term, the white noise power, was held fixed at a value of 2.0, as justified by earlier Exosat results in the same mode^{3,11}. The existence of QPOs in the highest intensity interval depends on this choice, that is, on the assumption that the source does not produce intrinsic white noise power. The second term, an exponential with integral A and e -folding frequency B , describes the LFN (power law fits to the LFN were not acceptable). The third term is a lorentzian peak profile with integral C , centroid frequency D , and half-width at half-maximum E . The results are summarized in Table 1. All errors are standard deviations (s.d.) ($\chi^2_{\min} + 1$) for one parameter of interest¹².

The spectral analysis was complete before the QPO discovery¹³. Energy spectra (1.6–20 keV) were derived for five intensity intervals (which are reasonably close to those used for

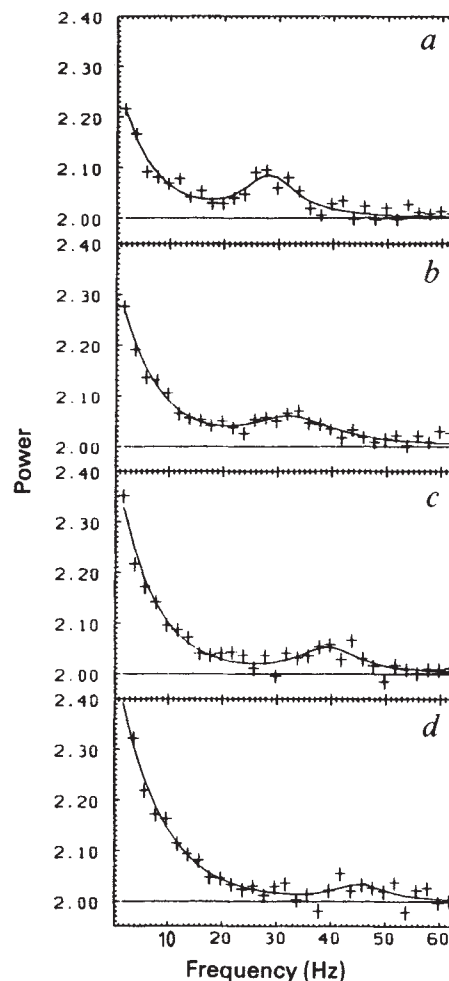


Fig. 1 Average power spectra of Cyg X-2 for four intensity intervals *a*, 800–850 c.p.s.; *b*, 850–900 c.p.s.; *c*, 900–1,000 c.p.s.; *d*, 1,000–1,200 c.p.s. Curved lines, functional least-square fits (see text). The power spectra would follow horizontal lines if there were no intrinsic source signal above counting statistics.

the time series analysis). Statistically acceptable fits require at least three simple 'standard' components: (1) a black-body spectrum ($kT \sim 1.2$ keV); (2) a thermal bremsstrahlung spectrum ($kT \sim 7$ keV); and (3) a very weak iron emission line. GSPC data fixed the energy, FWHM and equivalent width of the iron line at 6.5 ± 0.1 keV, 0.6 ± 0.2 keV, and 41 ± 10 eV, respectively. Spectra with fewer components and other combinations of three standard spectra, for example, including a 'multi-colour' disk black-body spectrum¹⁴, gave unacceptable fits. It is not obvious that there is a physical mechanism which could lead to the optically thin thermal bremsstrahlung spectrum observed in this case^{15–17}, but it fits the second component well. For the lack of a better term we will call it the 'thermal component'. Unsaturated componentized spectra for the thermal component^{13–18} did not improve the fits significantly, but yielded an optical depth of ~ 9 and an electron temperature of ~ 4 keV (using a solution to the Komaneets equation by Sunyaev and Titarchuk¹⁸). Figure 2 shows the correlation of luminosity with temperature for the black-body and the thermal component. Plotted errors are s.d. ($\chi^2 + 5.9$) for four parameters of interest¹². The black-body component is always weaker than the thermal one. The thermal component is stable within 20%, whereas the black-body luminosity varies by a factor of 2.7 without a change in its temperature.

We related the centroid frequency of the QPO to the luminosity of the individual spectral components. A power-law dependence $f \sim L^\alpha$ between frequency and black-body-, thermal-, and

§ Present address: The Max-Planck-Institut für Extraterrestrische Physik, 8046 Garching, FRG.

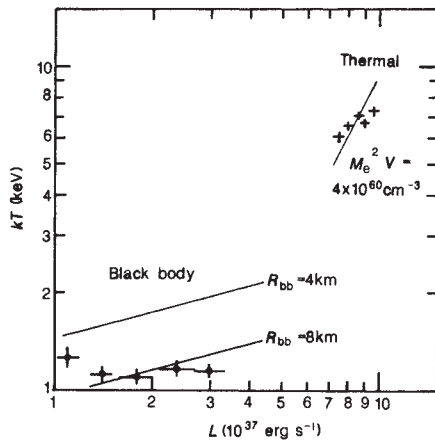


Fig. 2 X-ray luminosity (for a distance of 8 kpc (ref. 9), and assumed isotropic radiation) versus the temperature of the black-body and thermal components. Solid lines: lower left, tracks for constant black-body radii; upper right, the track of thermal bremsstrahlung with a constant emission measure. The luminosity of the thermal component remains approximately constant, but that of the black-body component changes by a factor of ~ 2.7 without a noticeable change in the black-body temperature. Consequently, the deduced black-body radius changes from ~ 6 to ~ 12 km.

total luminosity, yields $\alpha_{BB} = 0.59 \pm 0.05$, $\alpha_{Th} = 2.6 \pm 0.2$, and $\alpha_{tot} = 1.67 \pm 0.14$, respectively. Of course, other functional forms also fit the data.

QPOs with much longer periods and often with higher coherence, are common in cataclysmic variables (see for example, ref. 19). Although their detailed properties vary considerably even in the class of cataclysmic variables, some of them may share a similar origin with QPOs in the bright X-ray sources¹⁻⁶. We restrict our discussion to models which assume a neutron star with considerable magnetic field. We do not imply, however, that non-magnetospheric models (reviewed in ref. 3) can be excluded.

In a model by Bath²⁰, the oscillations represent the keplerian frequency of plasma clumps (plasmoids) at the magnetospheric boundary. For details see ref. 3; our calculations are identical to theirs, except that we used the magnetospheric radius for disk accretion²¹ instead of spherical accretion. For Cyg X-2 we find a magnetic field, B , of 5×10^{10} G, and a corresponding magnetospheric radius, r , of 150 km. The Van der Klis *et al.*³ data imply very similar values for GX 5-1. The change of frequency with total luminosity is $\alpha_{tot} = 1.7$, also comparable to that for GX 5-1 (ref. 3); however, both are larger than the expected α of 3/7 (refs 3, 20). To explain this discrepancy, Alpar and Shaham^{22,23} suggested that the observed frequency is a beat between the keplerian frequency and the neutron star rotation frequency (f_s). This idea was earlier suggested by Warner²⁴ for dwarf nova oscillations. For a harmonic number²⁵ $n = 1$ (one intensity maximum per beat cycle), we deduce $f_s = 108 \pm 15$ Hz, $B = (9.9 \pm 1.3) \times 10^9$ G, and $r \sim 55$ km. Similar values are obtained for GX 5-1 (ref. 3), where $f_s = 95.2 \pm 3.6$ Hz, $B = 1.8 \times 10^{10}$ G (disk accretion), and $r \sim 60$ km. Morfill and Trümper²⁶ suggest a beat-frequency model in which a rotating non-spherically symmetric Alfvén shell excites two plasma density waves interacting with the plasmoids. Here, the harmonic number²⁵ $n = 2$, and we find $B = (2.2 \pm 0.3) \times 10^{10}$ G and $f_s = 54 \pm 8$ Hz.

The relative amplitude (in terms of an r.m.s. equivalent sinusoidal modulation) of the low-frequency noise (LFN) is anticorrelated to that of the QPO. In GX 5-1, the QPO and LFN amplitudes are, however, proportional³. This is not the case here, and in Sco X-1, where the QPO can be strong and the LFN nearly absent⁶. In magnetospheric models invoking plasmoids^{20,22-26}, the LFN is a 'logical consequence' of QPO

and its amplitude is in first order expected to follow that of the QPO (however, see ref. 26). For any magnetospheric model it is disturbing that coherent pulsations are not detected. Reasons why they may not be observable (for example, intrinsically weak beaming which is further isotropized in a scattering cloud) are discussed^{3,25,27}.

In Cyg X-2 and GX 5-1 (refs 2, 3), there is a strong correlation between the frequency of QPO and the total source flux. Recent spectral studies^{13-17,28-33} show that many low-mass binaries (including Cyg X-2, GX 5-1, and Sco X-1) have two-component spectra. If the oscillations are of magnetospheric origin the mass accretion rate, $\dot{M}$, onto the magnetosphere (not the luminosity) determines the magnetospheric radius²¹ and the oscillation frequency. $\dot{M}$ cannot be unambiguously deduced from the spectra; the results are model dependent. In interpreting the spectra, it is commonly assumed that the disk reaches the neutron star¹³⁻¹⁷. The highly variable black-body component is then attributed to the neutron star—as suggested by the derived black-body radii of order 10 km. The more stable, and more luminous, thermal component is assumed to come from the inner disk. It has been suggested that large variations in the black body component can be produced by small changes in $\dot{M}$ (refs 14-18); in this case the observed black-body flux is not a good measure for $\dot{M}$, but the thermal flux would be. For Cyg X-2, the dependence of frequency on accretion rate ($\alpha' = d(\log f)/d(\log \dot{M})$) is then $\alpha_{Th} = 2.6$ (see above).

Lewin and Van Paradijs³⁴ have pointed out that the above interpretation may be incorrect for low-mass binaries with an evolved companion such as Cyg X-2, where a moderate neutron star magnetic field is expected. For a magnetospheric radius of the order of 10^2 km, any disk component could only contribute a small fraction ($\sim 10\%$) to the total luminosity. An interpretation of the spectra not in conflict with the magnetospheric models assigns the thermal component to the neutron star, and the black-body component to the inner disk. When $\dot{M}$ increases, the neutron star may not be able to accept this increase, if it already radiates close to the Eddington limit. The excess mass would be blown off as an accretion driven wind. This could explain why the thermal flux changes less than the black-body

Table 1 Parameters of the power spectra in four intensity intervals

Interval j	I	II	III	IV
$M(j)$	6,565	5,698	4,951	4,761
$I_{min}-I_{max}$ (c.p.s.)	800-850	850-900	900-1,000	1,000-1,200
I_{mean} (c.p.s.)	818	874	939	1,097
$A(P_{LFN})$	1.75 ± 0.11	2.33 ± 0.15	2.89 ± 0.12	3.96 ± 0.13
B (Hz)	6.3 ± 0.7	6.9 ± 0.6	6.9 ± 0.5	8.2 ± 0.5
$C(P_{QPO})$	1.58 ± 0.2	1.77 ± 0.35	1.13 ± 0.23	0.61 ± 0.20
D (Hz)	28.3 ± 0.6	32.4 ± 1.1	39.4 ± 1.2	45.3 ± 2.0
E (Hz)	6.0 ± 0.6	9.8 ± 1.3	6.9 ± 1.0	6.0 ± 1.8
$\chi^2/26d.f.$	1.27	0.71	1.14	1.22
a_{LFN} (%)	4.63 ± 0.15	5.16 ± 0.17	5.55 ± 0.12	6.01 ± 0.10
a_{QPO} (%)	4.8 ± 0.3	5.0 ± 0.5	4.1 ± 0.4	2.9 ± 0.5
β_{QPO}	0.847	0.800	0.723	0.648
L_{BB} (erg s ⁻¹)	1.2×10^{37}	1.6×10^{37}	1.9×10^{37}	2.8×10^{37}
L_{Th} (erg s ⁻¹)	8.0×10^{37}	8.2×10^{37}	8.7×10^{37}	9.7×10^{37}

Parameters of functional fits to the average power spectra in Fig. 1. $M(j)$ is the number of power spectra averaged in the j th intensity interval. I_{min} , I_{max} and I_{mean} are the minimum, maximum and mean count rate (including a background of about 80 c.p.s.). The parameters are: A , integrated power in the LFN; B , e-folding frequency; C , integrated power in the QPO; D , centroid frequency; E , peak HWHM. The reduced χ^2 (26 d.f.) is also given. a_{LFN} and a_{QPO} give the r.m.s. amplitudes of an equivalent sinusoidal modulation of the power in the LFN, and the QPO, respectively: $a_{QPO} = (P_{QPO}/I_{mean} * \beta_{QPO})^{0.5}$; $a_{LFN} = (P_{LFN}/I_{mean})^{0.5}$; β_{QPO} is a window function introduced by the bin size¹⁰. These amplitudes correspond to those given by Van der Klis *et al.*³ which should also be regarded as r.m.s. values. L_{BB} , and L_{Th} give the black body and the thermal luminosity (1-20 keV; for a distance of 8 kpc)⁹ respectively, interpolated from the spectral fits.

flux, which presumably follows the $\dot{M}$ changes. The appropriate value of α' is then $\alpha_{BB} = 0.59$, in reasonable agreement with the value of $1/3$ expected for radiation outside the magnetosphere²³. A magnetospheric scenario²⁰ without a beat mechanism could, therefore, fit Cyg X-2.

Other interpretations of the two-component spectra are possible (see refs 27, 28), however, all are very speculative. We can not judge whether either of the two spectral components in Cyg X-2 reflects the mass flux through the magnetospheric boundary layer. The association of spectrally resolved QPO with either of the two components would be a great help in understanding the QPO, and could also give some insight into the physics underlying the two-component spectra.

We thank Jan van Paradijs, Fred Lamb and Bill Friedhorsky for comments on various versions of this manuscript. W.H.G.L. acknowledges a generous award from the Alexander von Humboldt Stiftung and support from the John Simon Guggenheim Foundation.

Note added in proof: Since this paper was submitted, we have observed Cyg X-2 again several times and found QPO in all observations. In the meantime, four more QPO sources have been discovered, so that a total of seven objects are known now. For details see IAU Circulars 4081, 4101, 4102, 4110, 4116, 4117, 4140, 4147 and 4153.

Received 17 July; accepted 13 November 1985.

1. Hasinger, G., Langmeier, A., Sztajno, M. & White, N. E. *IAU Circ.* No. 4070 (1985).
2. Van der Klis, M. *et al.* *IAU Circ.* No. 4043 (1985).

3. Van der Klis, M. *et al.* *Nature* **316**, 225 (1985).
4. Middleditch, J. & Friedhorsky, W. *IAU Circ.* No. 4060 (1985).
5. Van der Klis, M., Jansen, F., White, N., Stella, L. & Peacock, A. *IAU Circ.* No. 4068 (1985).
6. Van der Klis, M. *Pap. at NATO Workshop on the Evolution of Galactic X-Ray Binaries*, Schloss Ringberg (1985).
7. Turner, M. J. L., Smith, A. & Zimmermann, H. U. *Space Sci. Rev.* **30**, 513-524 (1981).
8. Peacock, A. *et al.* *Space Sci. Rev.* **30**, 525-534 (1981).
9. Cowley, A. P., Crampton, D. & Hutchings, J. B. *Astrophys. J.* **231**, 539-550 (1979).
10. Leahy, D. A. *et al.* *Astrophys. J.* **266**, 160-170 (1983).
11. Langmeier, A., Sztajno, M. & Trümper, J. *Adv. Space Res.* **5**, 121 (1985).
12. Avni, Y. *Astrophys. J.* **210**, 642-646 (1976).
13. Hasinger, G., Pietsch, W., Sztajno, M. & White, N. E. *Pap. at ESA Workshop on Cataclysmic Variables and Low-mass X-ray Binaries*, Bamberg (1985).
14. Mitsuda, K. *et al.* *Publ. astr. Soc. Jap.* **36**, 741-759 (1984).
15. Hasinger, G., Langmeier, A., Pietsch, W. & Sztajno, M. *Space Sci. Rev.* **40**, 233-238 (1985).
16. White, N. E., Peacock, A. & Taylor, B. G. *Astrophys. J.* (submitted).
17. White, N. E. *et al.* *Mon. Not. R. astr. Soc.* (submitted) (1985).
18. Sunyaev, R. A. & Titarchuk, L. G. *Astr. Astrophys.* **86**, 121-138 (1980).
19. Cordova, F. A. & Mason, K. O. in *Accretion Driven Stellar X-ray Sources* (eds Lewin, W. H. G. & Van den Heuvel, E. P. J.) 147-187 (Cambridge University Press, 1983).
20. Bath, G. T. *Nature phys. Sci.* **246**, 84-87 (1973).
21. Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296-316 (1979).
22. Alpar, M. A. & Shaham, J. *IAU Circ.* No. 4046 (1985).
23. Alpar, M. A. & Shaham, J. *Nature* **316**, 239 (1985).
24. Warner, B. in *Cataclysmic Variables and related Objects* (eds Livio, M. & Shaviv, G.) 155-172 (Reidel, 1983).
25. Lamb, F. K., Shibasaki, N., Alpar, M. A. & Shaham, J. *Nature* (in the press).
26. Morfill, G. & Trümper, J. *Nature* (submitted).
27. Lamb, F. K. *Pap. at NATO Workshop on the Evolution of Galactic X-Ray Binaries*, Schloss Ringberg (1985).
28. Hirano, T. *et al.* *Publ. astr. Soc. Jap.* **36**, 769-784 (1984).
29. Kendziorra, E., Collmar, W., Brunner, H., Staubert, R. & Pietsch, W. *Space Sci. Rev.* **40**, 361-366 (1985).
30. Sztajno, M., Trümper, J., Hasinger, G. & Langmeier, A. *Space Sci. Rev.* **40**, 293-296 (1985).
31. Langmeier, A., Sztajno, M., Trümper, J. & Hasinger, G. *Space Sci. Rev.* **40**, 367-370 (1985).
32. Tanaka, Y. *Pap. at NATO Workshop on the Evolution of Galactic X-Ray Binaries*, Schloss Ringberg (1985).
33. Trümper, J. *et al.* *Space Sci. Rev.* **40**, 255-262 (1985).
34. Lewin, W. H. G. & van Paradijs, J. *Astr. Astrophys.* (in the press).

How old is Cygnus A?

M. J. Spinks, W. G. Rees & P. J. Duffett-Smith

Mullard Radio Astronomy Observatory, Cavendish Laboratory,
Madingley Road, Cambridge CB3 0HE, UK

We have mapped the extragalactic radio source Cygnus A at 81.5 MHz using a two-element interferometer with variable orientation and spacing (maximum 45 km). This is the lowest frequency at which the radio source has been mapped with sufficient resolution (16 arc s) to resolve the bridge of emission joining the two outer lobes, and our data provide new information about the radiation spectrum. By comparing our map with one made at 2.7 GHz, (ref. 1) we have measured the two-point spectral index as a function of distance along the source axis. We estimate the age of the radio source by using a model in which two oppositely directed beams ram their way out through the surrounding medium, depositing relativistic plasma as they do so^{2,3}. Synchrotron radiation losses are more severe at high frequencies, causing a break in the spectrum which occurs at progressively lower frequencies as the plasma ages, and our analysis suggests that Cyg A is ~3 Myr old.

Our interferometer, which will be described in detail elsewhere, consisted of part of a large fixed antenna (effective collecting area ~200 m²) at the observatory near Cambridge, together with a single Yagi antenna (~10 m²) which we moved around to obtain the desired coverage of the projected aperture plane (the $u-v$ plane). Tracking for a total of 6 h about meridian transit was possible, during which time the data were digitized (1-bit representation) at each antenna and recorded on video cassette tapes in eight contiguous frequency bands of 125 kHz each between 81 and 82 MHz. The signals were later processed using a purpose-built play-back terminal and correlator. We observed the radio source on each of 17 baselines for several consecutive days to check for repeatability and as an insurance against manmade interference. We found that the fringe amplitudes were reproducible to within 5% and that interference was never a serious problem, even when the portable array was set

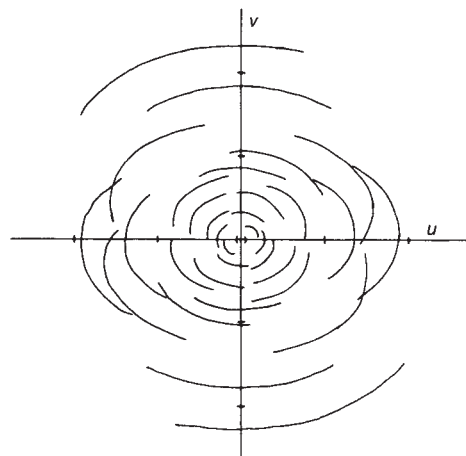


Fig. 1 Coverage of the $u-v$ plane. The pips are at intervals of $5,000\lambda$ (18.4 km).

up in the centre of Cambridge itself. The coverage of the $u-v$ plane is shown in Fig. 1.

The local oscillators at each antenna were locked to rubidium frequency standards, themselves locked to the ground-wave component of the 100-kHz LORAN-C signals from Sylt (Frisian Islands), which were sufficiently stable to allow coherent integration of the data for periods of at least 45 min. We used an integration time of 1 min so that our fringe amplitudes were not significantly affected by phase drifting. However, we were unable to correct the interferometer fringe phase for the residual random drifting of the local oscillators and ionospheric propagation effects. We therefore used a symmetric double-lobe model of Cyg A and assigned a phase of 0° or 180° to the fringe amplitudes, switching the phase wherever an amplitude passed through a zero. The shapes and magnitudes of the measured amplitudes near zeroes suggested that the antisymmetric component of the radio emission was <20% of the whole at 81.5 MHz, insufficient to cause appreciable error in the calculation of the ages.

The raw map, resulting from Fourier inversion of the measured amplitudes and assigned phases, was processed using

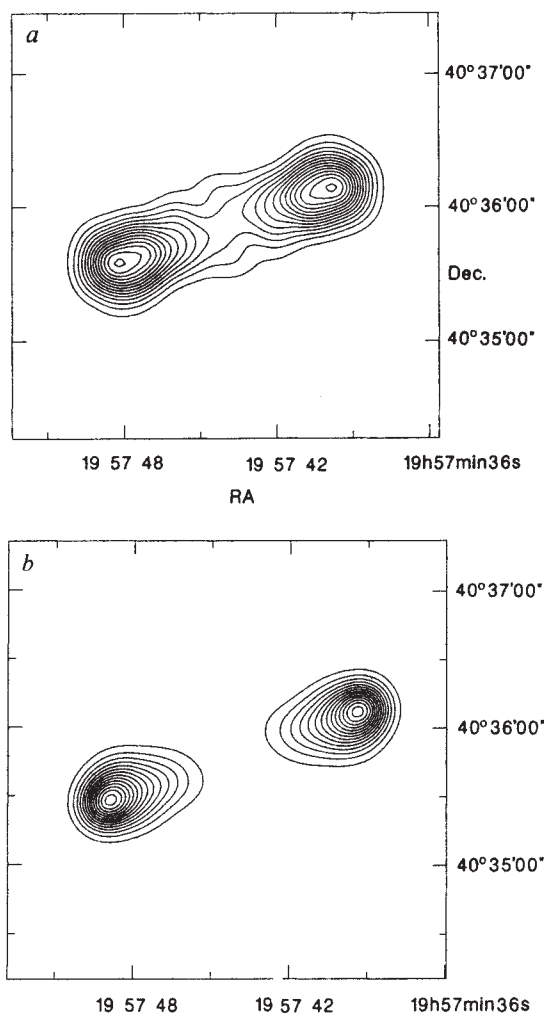


Fig. 2 *a*, The 81.5-MHz cleaned map. There are 14 contours at linear intervals of 150 Jy per beam, starting at 150 Jy per beam. *b*, The 2.7-GHz symmetric cleaned map convolved to the same resolution as in *a*. There are 14 contours at linear intervals of 25 Jy per beam, starting at 25 Jy per beam.

the CLEAN algorithm^{4,5}, and normalized to the flux density scale of Baars *et al.*⁶ (Fig. 2*a*). We believe it to have a dynamic range of more than 15:1, so that the lowest contour drawn in Fig. 2*a* is significant. A 2.7-GHz map made by Alexander *et al.*¹ using the 5-km telescope was processed in a similar fashion and then convolved to the resolution of the low-frequency map (Fig. 2*b*). As expected from the model, the spectral index steepens markedly towards the centre of the radio source.

The age of the plasma, t , depends on the characteristic synchrotron break frequency, ν_T , and the magnetic flux density, B , and is given by⁷

$$t = 1.06 \times 10^6 (\nu_T B^3)^{-0.5} \text{ s}$$

The magnetic flux density can be estimated by requiring that the total energy within the plasma is a minimum, giving nearly equal partition between the magnetic field and the particles. This leads to the formula⁷

$$B = 5.67 \times 10^{-9} \left[\frac{1+k}{\eta} (1+z)^{(3+\alpha)} \frac{F \nu^\alpha}{\theta \Phi S} \times \frac{\nu_2^{(1/2-\alpha)} - \nu_1^{(1/2-\alpha)}}{\frac{1}{2} - \alpha} \right]^{2/7} \text{ T}$$

where k (assumed equal to 1) is the ratio of proton to electron energies, η ($=1$) is the filling factor, F (Jy) is the flux density

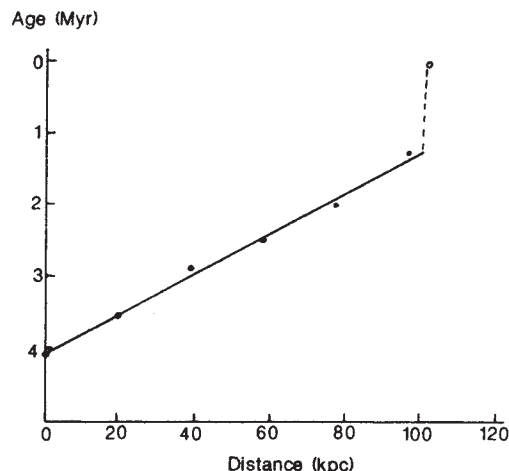


Fig. 3 Spectral age versus distance along the axis from the centre of the source. The open circle indicates the positions of the hotspots seen in the full-resolution 2.7-GHz map.

at frequency ν (GHz) arising from an area of the plasma θ by Φ arcs in extent, S (kpc) is the depth of the region in the line-of-sight, α is the spectral index (flux density proportional to $\nu^{-\alpha}$), ν_1 and ν_2 are the lower and upper frequencies defining the range in which there is significant synchrotron emission (we take $\nu_1 = 10$ MHz and $\nu_2 = 100$ GHz), and z ($=0.0574$)⁸ is the redshift. Our calculations are based on the further assumptions that the source is cylindrically symmetrical about its long axis and lies in the plane of the sky⁹, that the injection energy spectrum of the particles is constant in time with the value required to give the observed radiation spectral index (0.6) at low frequencies, and that the Hubble constant $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Note that the value of B is insensitive to errors in α because of the $2/7$ exponent in the above equation; choosing $\alpha = 0.7$ changes the age of the source by $\sim 10\%$.

Applying this analysis to our data gives a value of $B = 5.3 \text{ nT}$, with about 1 nT variation leading to the ages shown in Fig. 3. We deduce that Cyg A itself is ~ 3 Myr old, on the assumption that the abrupt change in slope at the points where the beams impinge upon the external medium (the 'hotspots') is caused by rapid ageing of the plasma in strong local magnetic fields. Our value is significantly lower than the only previous estimate which was based on a map made at 151 MHz by Winter *et al.*¹⁰, who obtained an age of 6 Myr with the same assumptions. However, their equivalent of Fig. 3 shows considerable curvature. When only the central part of their curve is considered, there is close agreement with our value. The slope of the line in Fig. 3 shows that the hotspots must be moving apart at $\sim 0.11c$ (where c is the velocity of light) at constant speed over most of the age of the source. The same speed was obtained by Alexander *et al.*¹ by observing spectral ageing within the outer lobes themselves.

The age of the Universe depends on the value of the Hubble constant and cosmological model adopted. Popular values lie in the range 10,000–18,000 Myr. The central galaxy in Cyg A must have been around for a substantial fraction of this time. It seems, therefore, that the phenomenon of double radio emission in this radio source must be comparatively short-lived. The galaxy either erupts once and dies (in radio terms), or else undergoes repeated outbursts. However, we find no evidence of emission from previous recent eruptions which would be expected to show up in a low-frequency map.

We thank many land owners around Cambridge who allowed us to erect the portable outstation and P. Alexander for discussions. M.J.S. acknowledges the support of a studentship from the SERC, and W.G.R. the support of a Research Fellowship at Christ's College, Cambridge.

Received 20 September; accepted 10 December 1985.

1. Alexander, P., Brown, M. T. & Scott, P. F. *Mon. Not. R. astr. Soc.* **209**, 851-868 (1984).
2. Rees, M. J. *Nature* **229**, 312-317 (1971).
3. Scheuer, P. A. G. *Mon. Not. R. astr. Soc.* **166**, 513-528 (1974).
4. Hogbom, J. A. *Astr. Astrophys. Suppl.* **15**, 417-426 (1974).
5. Clark, B. G. *Astr. Astrophys.* **89**, 377-378 (1980).
6. Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K. & Witzel, A. *Astr. Astrophys.* **61**, 99-106 (1977).
7. Pacholczyk, A. J. *Radio Astrophysics*, Ch. 6, 139-161 (Freeman, San Francisco, 1970).
8. Spinrad, H. & Stauffer, J. R. *Mon. Not. R. astr. Soc.* **200**, 153-158 (1982).
9. Hargrave, P. J. & Ryle, M. *Mon. Not. R. astr. Soc.* **166**, 305-327 (1974).
10. Winter, A. J. B. *et al. Mon. Not. R. astr. Soc.* **192**, 931-944 (1980).

Ionic origins of carbenes in space

Diethard K. Bohme

Department of Chemistry and Centre for Research in Experimental Space Science, York University, Downsview, Ontario, Canada M3J 1P3

Among the many molecules now known to enrich interstellar space are carbenes, molecules which have two electrons available on a carbon atom for chemical bonding. The dicarbenes :C=C: and :C=C=C: occur widely in the Universe, mainly in comets and the atmospheres of carbon stars¹. Other carbenes, including: :C=S , :C=O , :C=C=O and :C=NH , have been identified in cool molecular gas clouds¹. The origin of these molecules may involve ionizing radiation which, in these environments, can drive reactions between ions and molecules towards the formation of ionic carbene precursors. I now report experiments which suggest that a much higher proportion of the ion chemistry in partially ionized regions in space is directed towards the formation of carbene molecules than has previously been recognized. I identify plausible reaction paths which lead to both known and previously unknown carbenes, such as the cyclopropenylidene ring and long carbon-chain molecules. The cyclopropenylidene ring has recently been reported to be ubiquitous in the Galaxy^{2,3}, whereas long carbon-chain molecules have been implicated in the origin of diffuse interstellar lines⁴.

The reaction paths reported here that lead towards ionic carbene precursors have been identified using the selected-ion flow tube (SIFT) apparatus which allows the measurement of specific reaction rates and product distributions of individual ion/molecule reactions^{5,6}. Each step in a path towards ionic growth has been studied in isolation. The ion/molecule reactions reported here have all been observed to proceed near the collision rate at room temperature and hence without activation energy. This allows the laboratory results at room-temperature to be applied to the much cooler conditions of interstellar environments.

The atomic carbon ion is fundamental to the proposed carbene ion chemistry. Measurements in our laboratory have shown that C^+ reacts quickly with acetylenes and substituted acetylenes to establish carbene cations. With acetylene for example, the $\text{:C}_3\text{H}^+$ carbene cation, a resonance hybrid of $\text{HC}^+=\text{C}=\text{C:}$ and $\text{HC}\equiv\text{C}-\text{C:}^+$, is established at the collision rate with a reaction rate constant, k , of $2.2 \times 10^{-9} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ (ref. 7). Carbene ions of this type are most suited to molecular growth in two important ways. First they may grow heteroatomic chain molecules because of their ability to coordinate to non-bonded electron pairs according to reaction (1), where X is the heteroatom,



Second, they may grow hydrocarbon-chain molecules because of their tendency for C—H bond insertion according to reaction (2):

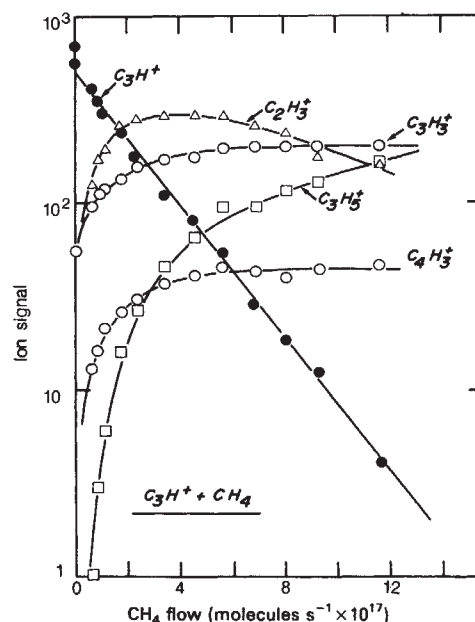
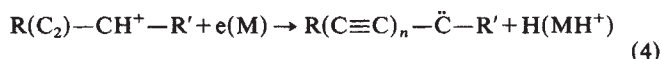


Fig. 1 Variations of ions observed with the addition of methane into the reaction region of a SIFT apparatus in which C_3H^+ has been established initially as the predominant ion⁸. The C_3H^+ is derived from propylene by electron impact at 45 eV. $P = 0.353$ torr, $\bar{v} = 7.4 \times 10^3 \text{ cm s}^{-1}$, $L = 46 \text{ cm}$, $T = 297 \text{ K}$, where P , $\bar{v}$, L and T are the pressure, average flow velocity, reaction length and temperature.

The product ion in reaction (2) may dissociate or lose energy by radiative or collisional stabilization. Reactions of types (1) and (2) will produce unsaturated neutral carbenes in subsequent neutralization reactions. These may involve H atom loss by recombination with an electron or the transfer of a proton to a recipient molecule, M . The neutral carbenes produced in this manner are indicated in reactions (3) and (4)



and involve either cumulative double bonds or conjugated triple bonds.

In reactions of type (1), $\text{:C}_3\text{H}^+$ has been shown to abstract an oxygen atom from various interstellar molecules to form an ion which may give tricarbon monoxide on neutralization⁸. There is also an analogous sulphur chemistry which leads to the prediction of the formation of tricarbon sulphide, :C=C=C=S , in the presence of H_2S in the interstellar medium. A higher multi-carbon monoxide ion was also observed to be formed in our laboratory by addition of :CO to $\text{:C}_3\text{H}^+$, which suggests an analogous addition by radiative association and thus formation of :C=C=C=O in interstellar regions rich in carbon monoxide. Of course, $\text{:C}_3\text{H}^+$ itself may neutralize by proton transfer or electron/ion recombination to form the dicarbene :C=C=C: .

Other reactions of C^+ are expected to lead to $\text{:C}_3\text{H}^+$ and higher members of the homologous $\text{H}(\text{C}_2)_n-\text{C:}^+$ series of carbene cations, as well as substituted carbene cations of the type $\text{R}(\text{C}_2)_n-\text{C:}^+$. For example, we have observed that C^+ reacts with cyanoacetylene to form $\text{:C}_3\text{H}^+$ and $\text{:C}_4\text{N}^+$ in a proportion of 4:1 with $k = 6.1 \times 10^{-9} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$, and with diacetylene to produce $\text{:C}_5\text{H}^+$ and $\text{:C}_3\text{H}^+$ in a proportion of 10:1 with $k = 1.6 \times 10^{-9} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ (refs 9, 10). Similar reactions with higher polyacetylenes and the known interstellar polycyanoacetylenes¹ should produce the corresponding higher carbene cations: $\text{:C}_5\text{H}^+$, $\text{:C}_7\text{H}^+ \dots (\text{C}_2)_n=\text{CH}^+$ and $\text{:C}_6\text{N}^+$, $\text{:C}_8\text{N}^+ \dots \text{C}^+-(\text{C}_2)_n-\text{CN}$. The former groups of ions in turn should spawn the cumulative dicarbenes $\text{:}(\text{C}=\text{C})_n=\text{C:}$ directly

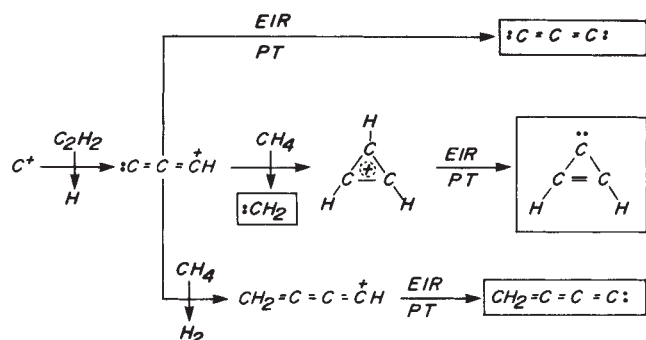


Fig. 2 Ion/molecule reactions which lead to the proposed formation of neutral carbenes from $:C_3H_3^+$. EIR, electron/ion recombination; PT, proton transfer.

by recombination with electrons or by proton transfer. Reactions with oxygen- or sulphur-bearing molecules will establish higher members of the unsaturated heteroatomic carbenes $:(C_2)_n=C=O$ and $:(C_2)_n=C=S$. The reactions of $:C^+-(C_2)_n-CN$ are uncertain but analogous reactions followed by neutralization by electron transfer could produce non-carbene chain radicals of the type $XC^+-(C_2)_n-CN$, where X may be O or S.

The reaction of $:C_3H_3^+$ with methane is particularly noteworthy as it allows formation of the cyclopropenylidene ring carbene in only three steps. This reaction has been observed to be rapid, $k = 5.5 \times 10^{-10} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$; $C_3H_3^+$ is one of the product ions (Fig. 1). Available thermochemical data dictate that only formation of the cyclopropenium ion is exothermic with methylene as the required neutral product⁸. Neutralization of this ion by electron/ion recombination or proton transfer can lead to cyclopropenylidene. We can therefore anticipate the formation of this ring carbene as well as methylene in ionized interstellar regions containing abundant acetylene and methane in the manner shown in Fig. 2. The formation of $C_3H_3^+$ from the reaction of $:C_3H_3^+$ with H_2 , which has been proposed as an alternative intermediate step in the formation of $:C_3H_2$, is less definitive because both the cyclopropenium and the linear propargyl cations are possible products¹¹. Figure 1 also demonstrates that another channel for the reaction of $:C_3H_3^+$ with CH_4 is the production of $C_4H_3^+$ which may neutralize to form the carbene $:C=C=C=CH_2$. The next member of the homologous series of these carbenes also appears to be accessible, in this case through the reaction of CH_3^+ with diacetylene, which has been seen to produce $C_3H_3^+$ in 1 of 10 reactive collisions with $k = 1.3 \times 10^{-9} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ (ref. 10). The $C_3H_3^+$ cation may neutralize to form $:C=C=C=CH_2$ (or $:CH-C\equiv C-C\equiv CH$).

Reactions of type (2) were observed to predominate, with the carbene cations $:C_3H_3^+$ and $:C_4H_3^+$ in reactions with HCN, CH_3CN , and HC_2CN . Exclusive addition reactions were observed in our laboratory to proceed with unexpectedly high efficiency at the collision rate^{8,9}. The high efficiencies suggest that analogous addition reactions proceed in the interstellar medium by radiative association as in reaction (5)¹².



Neutralization according to reaction (4) will then produce the carbene. Reactions of type (2) were observed which correspond to the possible formation of the substituted carbenes $:C(C_2H)CN$, $:C(C_2H)CH_2CN$, $:C(C_2H)C_2CN$ and $:C(C_2CN)_2$.

In summary, the observed ion chemistry clearly points towards the formation of many carbenes including methylene, completely unsaturated chain-like carbenes (with some incorporating an oxygen, sulphur or nitrogen atom), slightly hydrogenated chain-like carbenes and one aromatic ring carbene. I have shown that formation of these carbenes can be initiated by rapid reac-

tions of C^+ with simple interstellar molecules. Also, radiative association reactions are proposed which may lead to larger substituted carbene molecules. Formation of all these carbenes should provide further opportunities for molecular growth in the interstellar medium. This growth would not be restricted to reactions with ions but could involve reactions of carbenes with neutral molecules or other carbenes. In terrestrial organic chemistry, carbenes are among the most versatile and synthetically useful reactive intermediates, and they should have a similar function in extraterrestrial chemistry. For example, there are intriguing possibilities for the growth of graphite-like molecules from carbenes with either cumulative double bonds or conjugated triple bonds. Side-on interaction between two carbenes of one type or the other can lead to ring formation through cross-bonding, the inverse of the pyrolytic decomposition of graphite¹³. Alternatively, the long-chain carbenes may aggregate to form small interstellar grains, as suggested by Douglas⁴.

I thank the Natural Sciences and Engineering Research Council of Canada for financial support and Mr A. Fox for helpful discussions.

Received 28 August; accepted 22 November 1985.

1. Irvine, W. M. & Hjalmarson, A. in *Cosmochemistry and the Origin of Life* (ed. Ponnamperna, C.) 113-142 (Reidel, Dordrecht, 1983).
2. Vrtillek, J. M., Thaddeus, P. & Gottlieb, G. A. *Bull. Am. astr. Soc.* **17**, 568 (1985).
3. Matthews, H. E., Irvine, W. M., Madden, S. C. & Swade, D. A. *Bull. Am. astr. Soc.* **17**, 568 (1985).
4. Douglas, A. E. *Nature* **269**, 130-132 (1977).
5. Mackay, G. I., Vlachos, G. D., Bohme, D. K. & Schiff, H. I. *Int. J. Mass Spectrom. Ion Phys.* **36**, 259-270 (1980).
6. Adams, N. G. & Smith, D. J. *Phys.* **B9**, 1439-1451 (1976).
7. Bohme, D. K., Raksit, A. B. & Schiff, H. I. *Chem. Phys. Lett.* **93**, 592-597 (1982).
8. Raksit, A. B. & Bohme, D. K. *Int. J. Mass Spectrom. Ion Processes* **55**, 69-82 (1984).
9. Raksit, A. B. & Bohme, D. K. *Can. J. Chem.* **63**, 854-861 (1985).
10. Dheandhanoo, S., Forte, L., Fox, A. & Bohme, D. K. *Can. J. Chem.* (in the press).
11. Herbst, E., Adams, N. G. & Smith, D. *Astrophys. J.* **285**, 618-621 (1984).
12. Bates, D. R., *Astrophys. J.* **270**, 564-577 (1983).
13. Heimann, R. B., Kleiman, J. & Salansky, N. M. *Carbon* **22**, 147-156 (1984).

An intense cold core eddy in the North-East Atlantic

S. L. Kupferman*, G. A. Becker†, W. F. Simmons‡, U. Schauer†, M. G. Marietta* & H. Nies†

* Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

† Deutsches Hydrographisches Institut, Postfach 2 20, 2000 Hamburg 4, FRG

‡ Technadyne Engineering Consultants, Woods Hole, Massachusetts 02543, USA

We describe here a large intense mesoscale feature in the North Atlantic (48° N , 22° W) observed in April 1985. From the thermal structure of this feature at a depth of 475 m (Fig. 1a) and from the pattern of larger scale flow (dominated by a nearly rectilinear thermal front (NE-SW), Fig. 1b) we conclude that the feature is the eastern basin counterpart of the Gulf Stream cold-core eddies often observed in the western basin^{1,2}. To our knowledge, this is the most intense coherent cold-core eddy so far surveyed in this region, and the only such eddy for which data bearing on the mechanism of formation are available.

Western basin cold core eddies are spawned from Gulf Stream meanders that intensify, grow into loops, and pinch off from the main stream to form rings of current that wander independently in the waters to the south and south-east of the Gulf Stream³. They are readily distinguishable from other mesoscale features through their water mass properties, characterized by

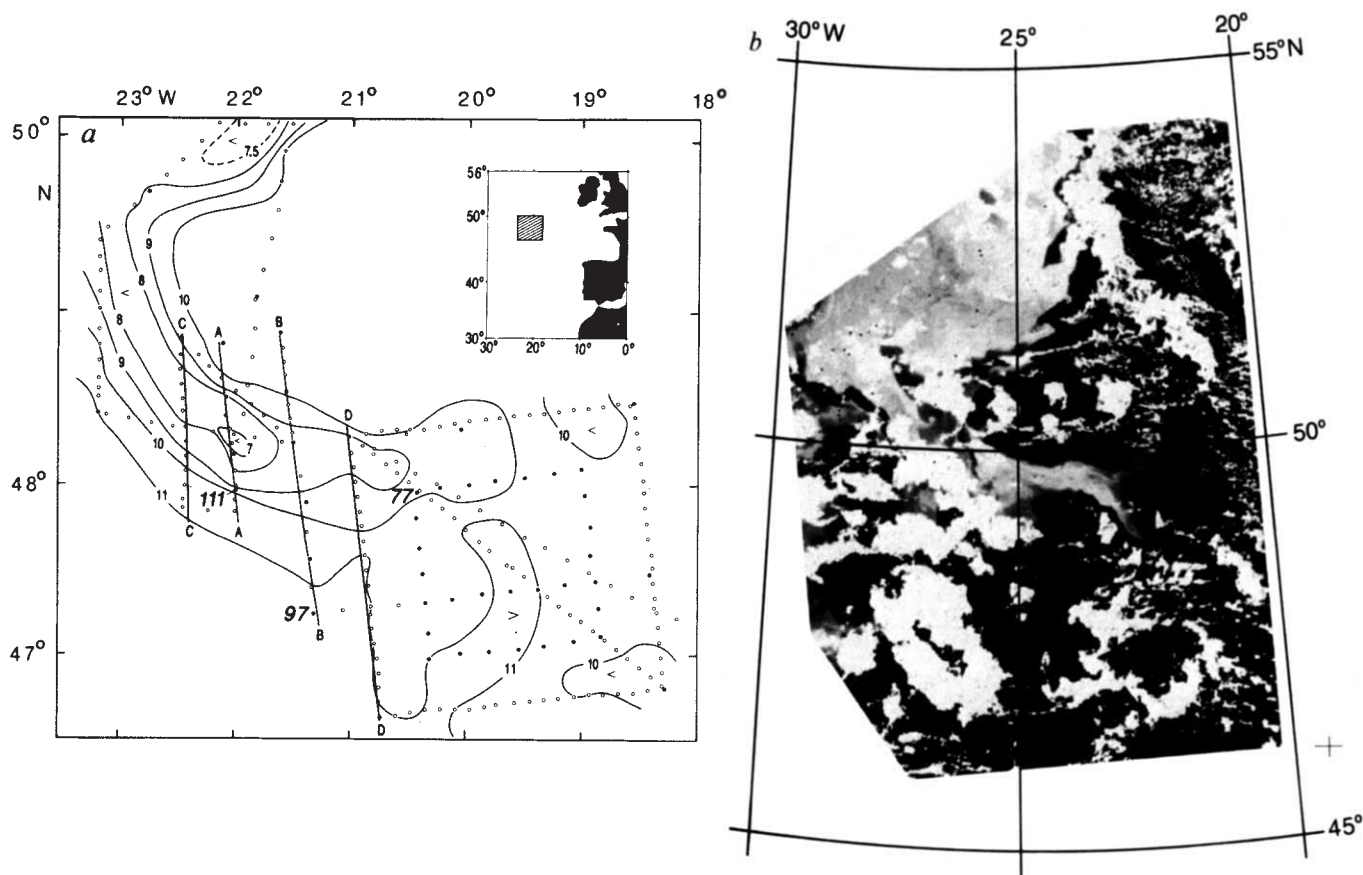


Fig. 1 *a*, Configuration of the cold-core eddy as defined by isotherm patterns at a depth of 475 m. The stations of the 185 km square and diagonal to the south-east were collected synoptically (3 days) at the outset of the cruise. They are primarily Sippican T-7 expendable bathythermographs (XBTs), plotted as open circles. The three north-south sections west of the square (A, B, C) were also synoptic and were occupied 10 days later. The irregular pentagon north and north-west of sections A, B and C was also synoptic and was occupied some 4 days after these sections. The 100-km box to the south-east (which is circumscribed by the 185-km box mentioned above) and the stations within it, were occupied during the 4 days after sections A, B and C were completed. They are mostly Neil Brown CTD stations, plotted as solid circles. *b*, Configuration of the surface temperature field in the neighbourhood of the eddy as defined by Advanced Very High Resolution Radiometric data from the NOAA 9 polar orbiter on 1 April 1985 (Ch. 4, 10.3–11.3 μm). Lighter areas are cooler, darker are warmer. Pure white areas are cloud tops. The NE-SW thermal front that dominates the field is identified as a downstream fragment of the Gulf Stream; that is, an element of the North Atlantic Current System¹². The feature corresponding to the eddy shown in *a* is roughly centre field, but is well to the north of the position at depth.

potential temperature (θ) and salinity (S). Water interior to the ring is cooler and fresher, like slope water. Water towards the exterior is warmer and saltier, like Sargasso Sea water³.

The eddy we observed is well-developed, but still attached to the parent current. Based on the 9°C isotherm at 475 m (Fig. 1*a*), the eddy is tongue-like in shape, nearly 300 km long (NW-SE) and as little as 40 km wide (NE-SW). The minimum observed temperature at 475 m is 6.8°C. It is found only 30 km south-west of the 10°C isotherm at the same depth.

The vertical structure of the eddy is shown in temperature section A-A of Fig. 2. We chose this section because the feature was intense and the data were synoptic within the limitations of shipboard expendable bathythermograph/conductivity-temperature-depth (XBT/CTD) work. The temperature section clearly shows isotherms rising and falling abruptly, characteristic of a cold-core eddy. The strongest gradients occur on the southern flank, where the isotherms are regularly spaced. The northern flank is somewhat gentler, with gradients strongest at depth and weakening toward the surface. Isotherm spacing is irregular. We will show that intrusive interleaving of different water masses occurs; it can cause such irregularities. The width of the feature, as measured by the 9°C isotherm at 475 m, appears to be 67 km, but as this section is skew to the eddy axis the actual width is closer to 47 km. The overall vertical isotherm excursion is about 400 m in the 8–11°C range. It is 450 m at

10.5°C. The surface temperature variability across the eddy is about 1.5°C (cooler in the interior). Therefore the eddy should be detectable by satellite radiometry given sufficient breaks in the clouds (Fig. 1*b*). Deep CTD data taken contemporaneously suggest the eddy structure persists to at least 2,500 m. The eddy signal may in fact penetrate to the bottom (4,400 m), but topography plays an important role in this region⁴ and topographically induced signals are strong enough over the bottom 1,000 m to mask eddy signals which are weaker at depth.

Geostrophic shears were calculated from deep CTD stations along section B-B of Fig. 1*a*. The feature there is less intense than in section A-A, but section B-B is longer and its CTD data are more plentiful and better spaced, which is important given the abrupt transitions in water mass properties. Geostrophic shears, with respect to 3,000 m, are strongest above 700 m. The relative velocity at the surface reaches a maximum of 36 cm s^{-1} towards the east along the southern flank of the meander. Considering the coarse resolution, maxima are probably higher. The maximum at 700 m is 19 cm s^{-1} , and at 2,700 m is a few cm s^{-1} . Transports are all towards the east over the southern part of the section, totalling 14 sverdrups (1 Sv = $10^6 \text{ m}^3 \text{ s}^{-1}$) above 1,200 m and 4 Sv between 1,200 m and 3,000 m for a total eastward transport of 18 Sv. Transports are to the west in the northern part of the section and are generally smaller. The maximum surface current there is 17 cm s^{-1} and the current

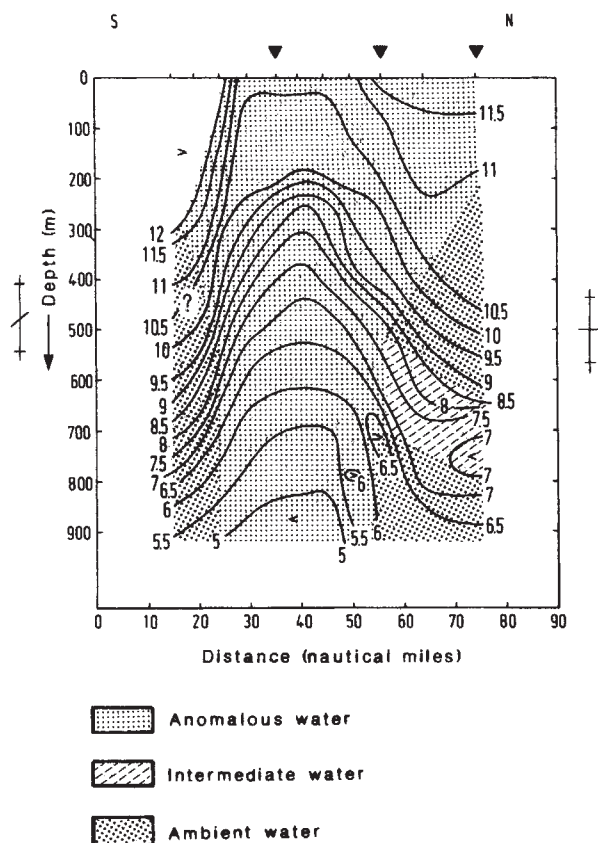


Fig. 2 A temperature section (A-A in Fig. 1a) through the eddy core. XBT stations are indicated by hatch-marks; three CTD stations are marked by triangles. The three principal water types, classified according to Fig. 3, are distinguished by stippling. The question mark on the southern flank connotes uncertainty due to the lack of CTD data.

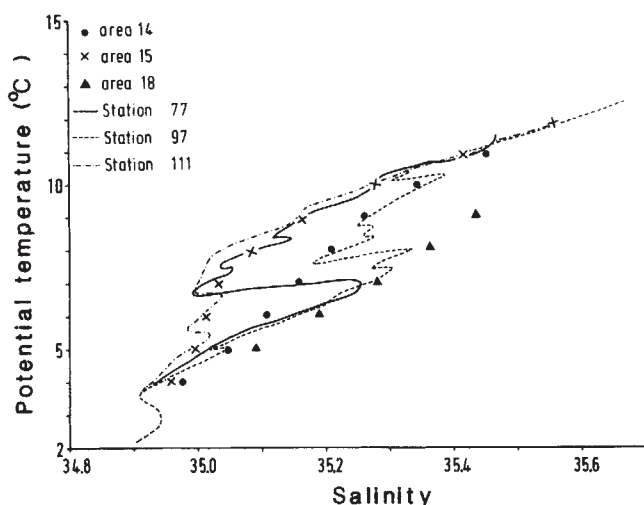


Fig. 3 θ - S diagram of three typical stations within the eddy and three reference θ - S relationships in the North Atlantic¹⁵. The three typical stations are: a, Station 97 (· · · · ·), characteristic of ambient type water; b, station 111 (— · — · —), characteristic of anomalous type water, and c, station 77 (—), characteristic of an abrupt shift between the two water types. The three reference stations are: d, Area 15 (x x x), in the nomenclature of Emery and Dewar¹⁵, which is typical of ambient conditions to the west (45° N, 32.5° W); e, area 14 (· · ·), typical of ambient conditions in the observed region¹⁵; f, area 18 (▲▲▲), typical of ambient conditions northeast of the observed region¹⁵.

at 700 m is 10 cm s^{-1} . Both of these values occur between the northernmost pair of stations. Above 1,200 m, the northernmost three stations show westward flow at 8 Sv whereas below 1,200 m only the northernmost two stations flow westward at 2 Sv. Thus, the eddy tends to lean somewhat to the south (see Fig. 2). The total westward transport observed on the northern flank is 10 Sv. Geostrophic shear calculations from other deep CTD stations give current patterns that are consistent with Fig. 1.

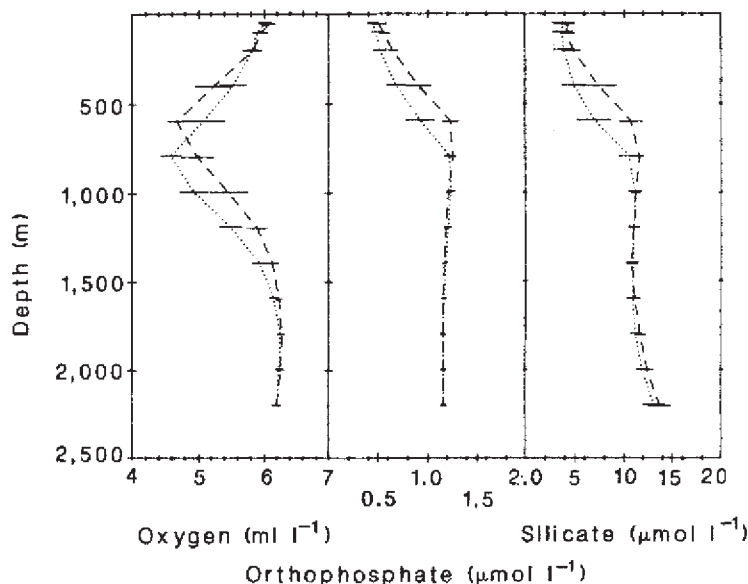
Current meter mooring data of 6 months duration 60 nautical miles to the south-east show brief bursts of speeds up to 20 cm s^{-1} at a depth of 377 m, with more frequent maxima of 10 cm s^{-1} (E. Mittelstaedt, unpublished data). An earlier current meter mooring of 8 months duration at 47° N, 15° W shows a month-long northward outbreak with speeds up to 22 cm s^{-1} at 400 m (ref. 5). It occurs in September, when the surface stratification, and therefore the width and range of influence of the average Subpolar Front, is at a maximum⁶. Thus, the instantaneous eddy kinetic energy associated with the cold-core feature reported here would have to be classified as generally higher than those previously reported⁷, and feasible only in the context of a narrow intense branch of the North Atlantic Current System (NACS). On longer timescales, the occurrence of such an eddy once per year could measurably increase the local annual mean eddy kinetic energy.

The idea of eddies in the NACS is not wholly new. Dietrich⁶ depicted two blister-like features in the summertime thermal structure of the Subpolar Front at 400 m just west of our survey. His station spacing, however, is not sufficiently dense to establish that the blisters are actually loops; that is, that water from the northwestern side has been drawn southeastwards and been encased in a circulating boundary of North Atlantic Central Water⁸ from the southern side of the front. Dickson and Gurbutt saw a similarly sized feature in a temperature section some 300 km to the north⁹ and Wegener reported a large-scale topographically induced feature to the west¹⁰; neither section included salinity. Howe and Tait¹¹ performed a CTD survey of the upper 500 m of a similarly sized feature about 600 km to the north in July. It too contained anomalous water. However, despite the stronger development of the seasonal thermocline at that time of year, their feature had no surface signature. Moreover, geostrophic shears increased with depth and were weaker than ours by a factor of roughly three. Using θ - S characteristics, they conjectured their feature had formed some 200 km to the west roughly 35 days before observation (based on an estimated migration speed of 0.08 m s^{-1}). A western basin eddy having these characteristics would have been thought of as older². They speculated that their eddy might be attached to a front, but the data were inconclusive. Using satellite-tracked surface drifters, Krauss and Käse¹² have shown the North-West Atlantic to have an abundance of mesoscale features. However, without hydrographic or geochemical data, it is impossible to distinguish features with water property anomalies from those without. In the western North Atlantic, the latter far outnumber the former¹³.

Figure 3 summarizes the observed water mass types and configurations based on the variation of their θ - S characteristics with depth; 'ambient water' is relatively warmer and has higher salinities due to the influence of Mediterranean Water. It is typical of the water in the North-East Atlantic^{5,14}. Below 1,200 m, ambient water is composed of Subpolar Intermediate Water (SIW). In contrast, 'anomalous water' has θ - S values characteristic of the Newfoundland basin^{14,15} to the west. Below 8 °C, anomalous water cools at roughly constant salinity ($\sim 35\text{‰}$), until it too joins the SIW below 1,200 m.

Many of our stations are either wholly ambient or wholly anomalous water. However, most involve both types. Typically, a station will start as one type, then will shift abruptly to the other. Often it will shift back again in an interleaving pattern. There is rarely evidence of intermediate water types. When intermediate water types are suggested, they are limited to thin

Fig. 4 Oxygen, orthophosphate and silicate profiles from the upper 13 of 24 samples taken throughout the water column. The values at each depth are averaged for ensembles of 39 ambient (---) and 23 anomalous (---) water stations as identified by θ - S characteristics (Fig. 3). The bars indicate one standard deviation.



bands of no more than 50–150 m in thickness. This, the relative intensity of the flow and its proximity to the rectilinear surface feature suggest that the feature is a newly forming cold-core eddy in the NACS.

Regions of ambient, anomalous and mixed water in the eddy are shown in Fig. 2. The pattern which emerges is that of a deep elongated lens of anomalous water riding along the centre and southern flank of the eddy feature, and surrounded on all sides by ambient water. We were unable to occupy a CTD station at the southern extreme of the section shown in Fig. 2, but other observations in the area during the same general time frame indicate the presence of ambient water to the south of the eddy.

XBT data from a March 1985 cruise in the same area show that the eddy feature was clearly present at that time on the section nearest to 21° W (D-D in Fig. 1b). It was still present and intense during a mid-May 1985 cruise, although its surface temperature signal had weakened. Its surface salinity signal remained.

By analysing pattern shifts between cruises, it is possible to infer that the feature as a whole moved to the north-east (~3 km per day) and that its principal axes rotated cyclonically (~0.5 deg per day). The eddy is better developed in mid-May in that more of the isotherms are closed (E. Mittelstaedt, personal communication). On-board efforts to map the eddy objectively in April show it to be attached to a small current fragment in the north-west corner of the computational domain. Preliminary dynamical forecasts¹⁶ lead to detachment, but the forecasted timescale for detachment is too short.

Analyses for levels of dissolved oxygen, orthophosphate and silicate were carried out at each station. Compared with ambient water, anomalous water showed larger but similarly distributed concentrations of phosphate and silicate from 0 to 800 m, with kinks at the oxygen minimum that are ~200 m shallower (Fig. 4). Oxygen profiles were related oppositely down to the oxygen minimum (~800 m). Below 800 m, anomalous water values are higher. Profile differences are discernible to roughly 1,600 m. Therefore, nutrient-depth profiles may be helpful in the identification of water mass anomalies. However, when nutrients and oxygen are plotted against temperature, water types can no longer be positively distinguished. Differences within water types are about equal to differences between types. This characteristic is also observed in western basin eddies¹⁷ and may

suggest that the decomposition of organic matter in the two water types is similar and possibly temperature-controlled.

There has recently been a sustained interest in the circulation and variability of the North-East Atlantic. It is an important region in terms of its unique flow characteristics, its position in the general circulation, its relationship to the European continent and the shallow northern seas, and to other applied and specific fields of study including fisheries and waste disposal research. We feel, therefore, that it is important to call attention to the existence of this intense eastern eddy feature with a view towards promoting awareness and stimulating additional hydrographic or other water-mass-sensitive observations in the region.

We acknowledge the support of the Bundesminister für Forschung und Technologie of the FRG and Sandia National Laboratories (SNL) through contract DE-AC04-76DP00789 from the US Department of Energy. We also thank the officers and crew of the RV *Meteor* for their support, the technicians and students of the Deutsches Hydrographisches Institut for analyses, and J. Garner of SNL for computer systems support. Satellite data were provided by M. Eckardt, image processing was performed by W. Tonn of the Freie Universität Berlin. E. Mittelstaedt discussed his May 1985 data with us before publication for which we thank him.

Received 8 July; accepted 30 October 1985.

- Richardson, P. L. *J. phys. Oceanogr.* **10**, 90–104 (1980).
- Joyce, T. & Weiler, P. *Oceans* **26**, 2, 34–44 (1983).
- Richardson, P. L. in *Eddies in Marine Science* (ed. Robinson, A. R.), 17–45 (Springer, New York, 1983).
- Becker, G. A., Bork, I. & Schauer, U. *Deutsches Hydrographisches Zeitung* **37** (Deutsches Hydrographisches Institut, Hamburg, 1984).
- Groupe Tourbillon Deep-Sea Res. **30**, 475–511 (1983).
- Dietrich, G. *Atlas of the Hydrography of the Northern North Atlantic Ocean* (Conseil Hydrographique Charlottenlund Slot, Copenhagen, 1969).
- Dickson, R. R., in *Eddies in Marine Science* (ed. Robinson, A. R.), 278–328 (Springer, New York, 1983).
- Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. *The Oceans* (Prentice-Hall, Englewood Cliffs, 1942).
- Dickson, R. & Gurburt, P. *Polymode News* **74**, 1–9 (Woods Hole Oceanographic Institution, 1980).
- Wegener, G. *Polymode News* **70**, 3–11 (Woods Hole Oceanographic Institution, 1979).
- Howe, M. R. & Tait, R. I. *Deep-Sea Res.* **14**, 373–378 (1967).
- Krauss, W. & Käse, R. H. *J. geophys. Res.* **89**, 3407–3415 (1984).
- Robinson, A. R. (ed.) *Eddies in Marine Science* (Springer, New York, 1983).
- Emery, W. J. & Dewar, J. S. *Prog. Oceanogr.* **11**, 219–305 (1982).
- Wright, W. R. & Worthington, L. V. *Am. Geogr. Soc. Folio* **19** (1970).
- Robinson, A. R. *et al. Nature* **309**, 781–783 (1984).
- Wiebe, P. H. *et al. Deep-Sea Res.* **23**, 695–710 (1976).

Nd–Sr–Pb isotope constraints on the sources of West Maui volcano, Hawaii

E. Hegner^{*†}, D. Unruh^{*} & M. Tatsumoto^{*}

^{*} US Geological Survey, PO Box 25046, Denver Federal Center, Denver, Colorado 80225, USA

[†] Colorado School of Mines, Golden, Colorado 80401, USA

The origin of the Emperor–Hawaiian volcanic chain is attributed to the northwesterly movement of the Pacific plate over a stationary mantle plume (hotspot)¹. There has been considerable controversy as to the nature and number of sources of Hawaiian hotspot volcanism. Thus far, most geochemical models have been based on rock suites that are not representative of fully developed volcanoes. Nd and Sr isotope ratios and trace element concentrations of volcanics from Haleakala (Maui), where all three volcanic stages are developed, have been interpreted as reflecting a mixing process of two isotopically distinct sources^{2,3}. In an attempt to test our earlier multiple-source model⁴, we have analysed Pb, Sr and Nd isotope ratios in volcanics from West Maui, the only other volcano with a complete volcanic record. Our results, presented here, indicate at least three isotopically distinct sources, one of which is heterogeneous with respect to Pb. Furthermore, the inferred depleted source for post-erosional volcanics has a Pb and Sr isotope composition intermediate between those of depleted and enriched mid-ocean ridge basalts (MORB, N-type and P-type), suggesting that this source is also heterogeneous.

Maui, located ~100 km north-west of Hawaii, comprises two volcanoes, Haleakala in the east and the smaller West Maui Volcano. All lavas are younger than 1.3 Myr and volcanism in the western part of Maui preceded that of Haleakala, which erupted last in AD 1790⁵. The evolution of fully developed Hawaiian volcanoes can be described by three main phases⁶: (1) the shield-building stage with voluminous tholeiitic basalts; (2) the caldera-filling or late-stage volcanism with both tholeiitic and alkalic rocks; and (3) the post-erosional stage with minor silica-undersaturated rocks (for example, basanites and nephelinites).

Pb, Sr and Nd isotope ratios for our samples from West Maui and Haleakala are listed in Table 1 and are shown in Figs 1–3. Sr and Nd isotope ratios vary inversely for lavas from West Maui, as do those from Haleakala and other Hawaiian volcanoes (Fig. 1). Shield-building tholeiites and caldera-filling alkalic lavas from West Maui volcano alone correlate negatively and plot at the upper end of the Hawaiian data array with indistinguishable Nd and Sr isotope ratios. A close genetic relationship of tholeiites and associated alkali basalts is also indicated for Waianae (Oahu)⁴ and Mauna Kea (Hawaii)^{7,8}. However, the isotopic dissimilarity of tholeiites and late-stage alkali basalts as seen at Haleakala², East Molokai⁹ and Kohala¹⁰ indicates a complex genetic relationship between these rock types. Where tholeiites are isotopically distinct from late-stage alkali basalts, they tend to have more enriched Nd and Sr compositions, as is well documented at Haleakala². We regard this as reflecting decreased lithosphere contamination of the tholeiitic magmas due to earlier established conduit systems and thinned lithosphere caused by nearby volcanism. On Hawaii, with five closely spaced volcanoes, the youngest tholeiites are not more enriched isotopically than the older ones. ²⁰⁶Pb/²⁰⁴Pb ratios, however, increase progressively along the Kea lineament⁷ from Kohala through Mauna Kea to Kilauea (Fig. 3). This indicates, in our opinion, increasing involvement of radiogenic Pb from old uppermost lithosphere (Kea component⁷) in the magma generation process and ongoing lithospheric thinning.

Post-erosional lavas from Maui have similar ¹⁴³Nd/¹⁴⁴Nd but lower ⁸⁷Sr/⁸⁶Sr ratios than the older volcanics, and they plot

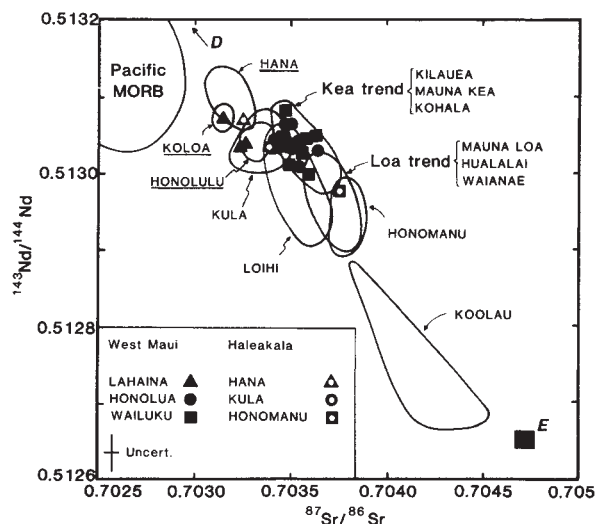


Fig. 1 Plot of ⁸⁷Sr/⁸⁶Sr versus ¹⁴³Nd/¹⁴⁴Nd, showing the data from West Maui and other Hawaiian volcanic rocks. Data points for shield-building (Wailuku) and caldera-filling (Honolua) rocks from West Maui are indistinguishable and plot with low ⁸⁷Sr/⁸⁶Sr ratios at the upper end of the Hawaiian trend. Post-erosional volcanics (Lahaina) from West Maui possess distinctly lower ⁸⁷Sr/⁸⁶Sr ratios than the associated volcanics and plot like other Hawaiian post-erosional suites (Hana, Kohala, Honolua, underlined) close to the Pacific MORB field. Also shown are the data for three samples from East Maui (Haleakala); symbols represent same volcanic stages as for West Maui. Inferred enriched (E) and depleted (D) sources are indicated. Data sources: refs 2, 4, 11, 14 and 16.

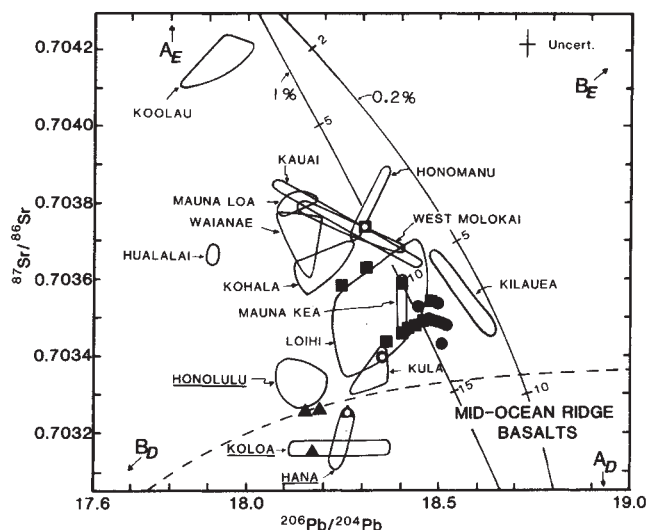


Fig. 2 Plot of ²⁰⁶Pb/²⁰⁴Pb versus ⁸⁷Sr/⁸⁶Sr, showing the data of shield-building and caldera-filling samples from West Maui plotting close to the Kilauea apex. The younger post-erosional volcanics have less radiogenic Pb and Sr compositions, similar to other post-erosional rocks, and plot near or in the MORB field. Endmembers are indicated by arrows. Subscripts E and D refer to enriched and depleted sources, respectively. Mixing lines for melts produced by 1% (tholeiites) and 0.2% (post-erosional rocks) melting of a depleted source and 13% and 14% melting of an enriched source are shown. Ticks on mixing lines indicate the proportion of melt from the depleted source. Modelling parameters for tholeiites and post-erosional rocks, Sr concentrations in endmembers and mineral/melt partition coefficients were taken from ref. 2. The Pb mineral/melt partition coefficient was inferred to be 0.001 for all phases. Pb concentrations from ref. 13; the Pb concentration of the enriched source is that of 'mean mantle', which is compositionally similar to the enriched source used in the Haleakala model². Data sources: refs 4, 11, 14, 24; MORB field from ref. 25.

Table 1 Isotope data for Maui volcanics

West Maui	$^{143}\text{Nd}/^{144}\text{Nd}$	$^{87}\text{Sr}/^{86}\text{Sr}$	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
Lahaina volcanics*					
C-130 basanite	0.51309 $\pm$ 3 0.51311 $\pm$ 1	0.70315 $\pm$ 3	18.169	15.452	37.750
C-115 basanite	0.51303 $\pm$ 2	0.70326 $\pm$ 2	18.191	15.460	37.833
HMT 79-5 (C-115)	0.51303 $\pm$ 2	0.70325 $\pm$ 4 0.70327 $\pm$ 3	18.155	15.432	37.735
Honolua volcanics					
C-155 mugearite	0.51304 $\pm$ 1	0.70343 $\pm$ 4	18.513	15.494	38.050
C-153 hawaiite	0.51304 $\pm$ 1	0.70353 $\pm$ 3	18.447	15.456	37.917
C-150 mugearite	0.51307 $\pm$ 1	0.70349 $\pm$ 3	18.502	15.483	38.033
C-149 mugearite	0.51304 $\pm$ 6	0.70349 $\pm$ 3	18.509	15.489	38.030
C-128 trachyte	0.51303 $\pm$ 2	0.70347 $\pm$ 3	18.492	15.476	38.008
C-116 trachyte	0.51306 $\pm$ 1	0.70348 $\pm$ 5	18.523	15.471	37.961
C-92 trachyte	0.51304 $\pm$ 2	0.70349 $\pm$ 3	18.482	15.466	37.956
HMT 79-1a (C-92)	0.51303 $\pm$ 2	0.70355 $\pm$ 4	18.487	15.462	37.947
HMT 79-2b alkali basalt†	0.51302 $\pm$ 3	0.70354 $\pm$ 3	18.506	15.560	38.040
Wailuku volcanics					
C-119 tholeiite	0.51304 $\pm$ 3	0.70347 $\pm$ 4	18.426	15.472	37.904
HMT 79-2a (C-119)	0.51303 $\pm$ 2	0.70356 $\pm$ 3	18.462	15.454	37.899
C-118 alkali basalt	0.51305 $\pm$ 1	0.70364 $\pm$ 4 0.70359 $\pm$ 2‡	18.258	15.471	37.762
C-117 tholeiite	0.51301 $\pm$ 2	0.70359 $\pm$ 3 0.70361 $\pm$ 2	18.404	15.492	37.940
C-114 alkali basalt	0.51303 $\pm$ 2	0.70346 $\pm$ 4	18.412	15.472	37.906
C-110 tholeiite	0.51305 $\pm$ 1	0.70363 $\pm$ 3 0.70365 $\pm$ 3‡	18.309	15.476	37.874
C-107 tholeiite†	0.51307 $\pm$ 1	0.70348 $\pm$ 3	18.438	15.499	37.948
C-100 tholeiite	0.51303 $\pm$ 1	0.70361 $\pm$ 4 0.70344 $\pm$ 3‡	18.365	15.470	37.877
East Maui (Haleakala)					
Hana volcanics					
C-129 basanite	0.51307 $\pm$ 1	0.70325 $\pm$ 3	18.259	15.453	37.828
Kula volcanics					
C-127 hawaiite	0.51303 $\pm$ 2	0.70339 $\pm$ 4	18.351	15.468	37.938
Honomanu volcanics					
C-122 tholeiite	0.51298 $\pm$ 2	0.70374 $\pm$ 3 0.70374 $\pm$ 4	18.303	15.452	37.924

Uncertainties are 2σ of the mean based on in-run statistics; repeated analyses of standards indicate an error of $\sim 2 \times 10^{-5}$ in the $^{143}\text{Nd}/^{144}\text{Nd}$ ratio for reproduction of data. Ratios are normalized to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$ and $^{86}\text{Sr}/^{88}\text{Sr} = 0.1194$. The following standard isotope ratios were measured: $^{143}\text{Nd}/^{144}\text{Nd}$, 0.511856 $\pm$ 10 (La Jolla Nd Standard) and 0.512654 $\pm$ 20 (BCR-1); $^{87}\text{Sr}/^{86}\text{Sr}$ 0.710265 $\pm$ 15 (NBS Standard SRM-987), 0.705029 ($n=2$, BCR-1). Pb isotope ratios are corrected for mass fractionation of 0.12 $\pm$ 0.03% per AMU based on analyses of NBS standard SRM-982 and are precise to 0.04% per AMU.

* Formerly designated 'Volcanic Series'. Sample numbers in parentheses indicate that rock types and sampling locations of HMT samples are the same as those of the corresponding C-sample. Sample locations are: C-149, Pohakupue Gulch; C-150, Honolulu Gulch; C-153, rock boulder from Kanaha Stream below Paupau Hill; C-155, Makamakaole Gulch; the remaining samples are from MacDonald's collection⁵.

† Analyses from P. Stille *et al.* (personal communication).

‡ Leached samples (6M HCl for 10 h); more than one analysis indicates re-determined values.

close to the MORB field (Fig. 1), as do other Hawaiian post-erosional volcanics. Chen and Frey^{2,3} found that Sr and Nd isotope ratios and the concentrations of some incompatible elements could be reproduced by variable mixing of melts generated from a depleted (MORB-like) and an enriched mantle source compositionally similar to inferred primitive mantle^{11,12}. The high incompatible-element concentrations of the alkalic and post-erosional volcanics could then be produced by a very low degree of melting of the MORB-like source. The Sr and Pb isotope data from West Maui require more than two isotopically distinct sources, as will be shown.

The $^{206}\text{Pb}/^{204}\text{Pb}$ - $^{87}\text{Sr}/^{86}\text{Sr}$ data for Hawaiian volcanics can be modelled with the two sets of endmembers shown in Fig. 2. Both endmember pairs yield analogous results in binary mixing calculations as established for Haleakala. Using endmember set A, one can see in Fig. 2 that the data points of shield-building and caldera-filling volcanics from West Maui and other Hawaiian islands require a swarm of mixing lines. Pb and Sr data of shield-building rocks from West Maui can be explained by 9–13% mixing of a melt derived from a MORB-like source with a melt from a reservoir with a 'mean mantle' composition¹³, once Sr and Pb concentrations are adjusted for crystal fractionation. Mixing calculations using the Haleakala model parameters for post-erosional lavas indicate that the compositions of the latter lie on mixing lines to the right of those for tholeiites and late-stage alkali basalts, in contrast to the situation depicted in Fig. 2; this results from the higher incompatibility of Pb

compared with Sr in mantle phases^{14,15}. Lowering the degree of partial melting of the MORB-like source enriches the melt in MORB-Pb relative to MORB-Sr. Consequently, the melt mixtures have a pronounced MORB-Pb signature and a large variation in their $^{87}\text{Sr}/^{86}\text{Sr}$ ratios (Fig. 2). This evidence requires contributions from more than two isotopically distinct sources and requires a more complex magma generation model than that proposed for Haleakala post-erosional lavas. It has been suggested that Hawaiian post-erosional lavas, with their relatively nonradiogenic Pb, could have originated in the uppermost lithosphere¹⁶. After ~ 80 Myr (the age of the oceanic crust underlying Hawaii) a MORB-like $^{238}\text{U}/^{204}\text{Pb}$ ratio of 20–30 would cause a $^{206}\text{Pb}/^{204}\text{Pb}$ ratio of 18.1 to evolve to 18.6. The relatively little radiogenic Pb in post-erosional lavas thus argues against an origin in the uppermost lithosphere.

Figure 3 shows the Pb isotope ratios of West Maui and other Hawaiian volcanics. The data show a wide range and plot on the Hawaiian trend. The apparent Pb isochron with an age of ~ 950 Myr⁷ must be interpreted as a mixing line because at least three components in Hawaiian volcanics are indicated. This conclusion is also supported by the Pb data in Fig. 3, which would lie on a straight mixing line in a two-endmember model. The highly fractionated caldera-filling rocks possess higher $^{206}\text{Pb}/^{204}\text{Pb}$ ratios than the shield-building volcanics. This is interpreted as reflecting contamination of the magmas by old oceanic crust with an evolved $^{206}\text{Pb}/^{204}\text{Pb}$ ratio during fractionation in shallow magma chambers.

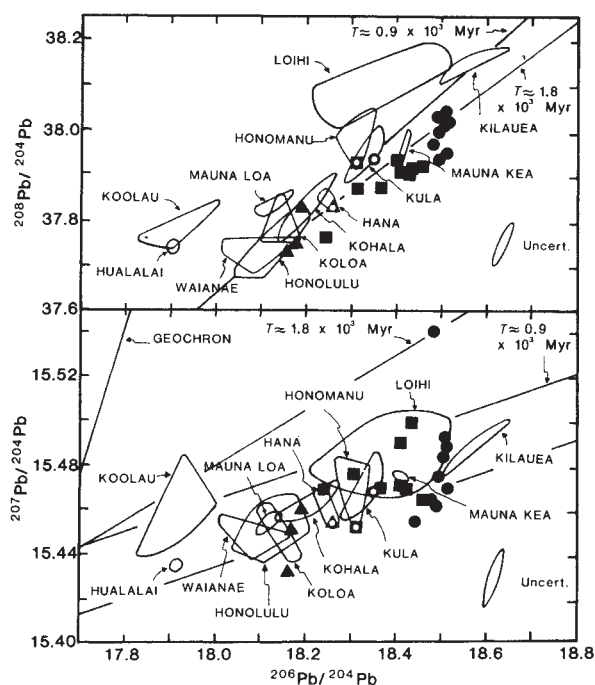


Fig. 3 Relationship of $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ to $^{206}\text{Pb}/^{204}\text{Pb}$ for Hawaiian lavas. Shield-building and caldera-filling volcanics from West Maui plot on the more radiogenic side of the Hawaiian trend. Fractionated caldera-filling rocks are slightly more radiogenic in Pb, while post-erosional lavas are less radiogenic in their Pb isotope composition than shield-building volcanics. T , Age calculated from the respective lines. Data sources: refs 4, 7, 14, 24.

Three major isotopically distinct sources are required to explain the sources of Hawaiian volcanism: Honolulu-Lahaina, Kilauea and Koolau, although the composition of the pure component will lie beyond the named endmembers (Fig. 2). Figures 1 and 2 show that Honolulu and Lahaina and other Hawaiian post-erosional volcanics plot with comparatively non-radiogenic Pb composition close to the MORB field, suggesting a depleted reservoir as their principal source^{6,14,17}. The Pb-Sr relationship indicates that the depleted component in the post-erosional rocks is a mixture of N- and P-type MORB. This delineates the asthenosphere as the probable source region, as has been proposed on the basis of petrological evidence from xenoliths in post-erosional lavas¹⁸.

Kilauea, the second component, is characterized by the most radiogenic Pb and moderately radiogenic Nd. Seismic data reveal that Kilauea and Mauna Loa share a common plumbing system at a depth of 30–50 km¹⁹. Therefore, their different Pb isotope compositions were probably created in the uppermost lithosphere or crust, our preferred Kilauea source.

Koolau data indicate a component less depleted than the other endmembers, with Sr and Nd ratios close to inferred primitive mantle values^{11,12}, and low $^{206}\text{Pb}/^{204}\text{Pb}$ ratios. On this basis we assign this component to the mantle plume fed by a primitive reservoir. He, Ar and Xe isotope ratios indicate the presence in some Hawaiian magmas of a component derived from a long-lived undegassed reservoir^{20–22}. Similar analyses on Koolau volcanics would be a useful further test of our hypothesis. The fact that the Koolau component is least represented in Hawaiian volcanics also agrees with both He and Ar isotope evidence indicating a proportion of less than 30% of plume component in the magmas from Kilauea²³.

Our results demonstrate the need for comprehensive geochemical studies on systematically sampled rock suites to elucidate further the magma generation at the Hawaiian hotspot.

We thank P. Lipman and Z. Peterman for commenting on an earlier version of the manuscript, and F. A. Frey for his detailed and constructive criticism.

Received 24 July; accepted 26 November 1985.

1. Wilson, J. T. *Can. J. Phys.* **41**, 863–870 (1963).
2. Chen, C. Y. & Frey, F. A. *Nature* **302**, 785–789 (1983).
3. Chen, C. Y. & Frey, F. A. *J. geophys. Res.* **90**(B10), 8743–8768 (1985).
4. Stille, P., Unruh, D. & Tatsumoto, M. *Nature* **304**, 25–29 (1983).
5. McDougall, I. *Bull. geol. Soc. Am.* **75**, 107–128 (1964).
6. Macdonald, G. A. & Katsura, T. *J. Petrol.* **5**, 82–133 (1964).
7. Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63–87 (1978).
8. Frey, F. A. *et al. EOS* **65**, 300 (1984).
9. Clague, D. H., Frey, F. A. & Beeson, M. H. *EOS* **64**(45), 902 (1983).
10. Hofmann, A. W., Raczek, I. & Feigenson, M. D. *Terra Cognita* **4**(2), 189 (1984).
11. O'Nions, R. K., Hamilton, P. J. & Evenson, N. M. *Earth planet. Sci. Lett.* **39**, 13–22 (1977).
12. Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7411–7422 (1979).
13. Davies, G. F. *J. geophys. Res.* **89**(37), 6017–6040 (1984).
14. Sun, S. S. *Phil. Trans. R. Soc. A* **297**, 409–445 (1980).
15. Watson, E. B., Ben Othman, D., Luck, J. M. & Hofmann, A. W. *Terra Cognita* **5**(1), 276 (1985).
16. Staudigel, H. *et al. Earth planet. Sci. Lett.* **69**, 13–29 (1984).
17. Feigenson, M. D., Knorr, R. & Raczek, I. *EOS* **65**(16), 301 (1984).
18. Sen, G. *Earth planet. Sci. Lett.* **66**, 215–228 (1983).
19. Klein, F. W. *J. geophys. Res.* **87**, 7719–7723 (1982).
20. Kaneoka, I. & Takaoka, N. *Earth planet. Sci. Lett.* **39**, 382–386 (1978).
21. Kurz, M. D., Jenkin, W. J. & Hart, S. R. *Nature* **297**, 43–47 (1982).
22. Allegre, C. J., Staudacher, T., Sarda, P. & Kurz, M. *Nature* **303**, 762–766 (1983).
23. Kaneoka, I. *Nature* **302**, 698–700 (1983).
24. Unruh, D., Stille, P. & Tatsumoto, M. *EOS* **64**(18), 347 (1983).
25. White, W. M. *Geology* **13**, 115–118 (1985).

Carbon and sulphur isotopic composition of Kilauea parental magma

Terrence M. Gerlach

Geochemistry Division, 1543, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

Donald M. Thomas

Hawaii Institute of Geophysics, University of Hawaii, 2525 Correa Road, Honolulu, Hawaii 96822, USA

All the available data for the isotopic composition of carbon in gases emitted from summit fumaroles at Kilauea volcano, Hawaii, indicate that the carbon is enriched in ^{13}C compared with oceanic tholeiite and in contrast to most estimates for the isotopic composition of mantle carbon^{1–6}. Several mechanisms proposed to explain this enrichment of ^{13}C have proved unsatisfactory². We show here that recent models^{7,8} for the volatile contents and degassing of Kilauea magma permit a satisfactory reinterpretation of carbon and sulphur isotope data^{1–3,9,10} for Kilauea. The results indicate that Kilauea parental magma is significantly enriched in ^{13}C compared with mid-oceanic basalts and with inferred isotopic compositions for mantle carbon^{3–6}. Hence, mantle sources of hot-spot and mid-oceanic basalts may differ in $\delta^{13}\text{C}$. In contrast, the isotopic composition of sulphur in Kilauea parental magma is similar to that of mid-oceanic basalts¹¹.

Kilauea volcano is detached from the south-east shield of Mauna Loa volcano by east and south-west rifts (Fig. 1)¹². The roof of the summit magma chamber is approximately 2 km under the caldera and 1 km under Halemaumau pit crater¹³. Mantle sources at 30–90 km depth supply parental magma to the chamber^{14–16}. The chamber has been described as a dyke-and-sill complex with a floor at 6 km, a crown at 1–2 km, and an east-rift duct at 1–2 km connected to the east-rift magma reservoirs^{12,17,18}. Communication between the summit and south-west rift reservoirs is poorly defined^{19,20}. A relatively steady supply of magma inflates the summit chamber before being discharged to the surface or into the rifts^{20,21}.

According to a recent model⁷, the melt fraction of the parental magma contains approximately 0.30 wt% H_2O , 0.65 wt% CO_2 and 0.13 wt% S. Because parental melt saturates in CO_2 long before reaching the summit chamber, the processes of reservoir equilibration during residence at the summit include separation of excess volatiles in a chamber gas containing about 90 vol.% CO_2 , 10 vol.% SO_2 and minor H_2O . Chamber-gas venting of

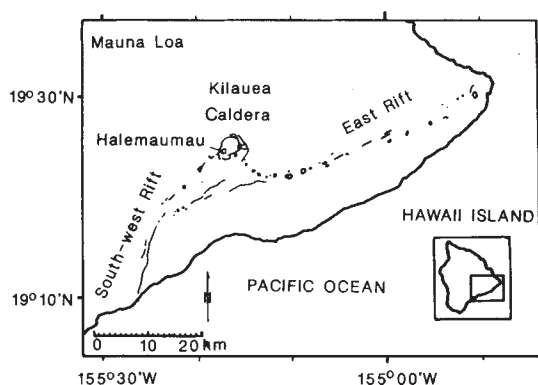


Fig. 1 Sketch map of Kilauea volcano, Hawaii. The east rift extends along its subaerial north-east trend for ~70 km under the sea.

CO₂ occurs at the impressive rate of 3,600 tons per day. The reservoir-equilibrated melt contains approximately 0.27 wt% H₂O, 0.07 wt% S and 0.02–0.10 wt% CO₂ (depending on the depth of equilibration). Thus, most of the parental CO₂ is degassed from the summit chamber. An alternative model⁸, derived from different geochemical data, proposes similar volatile contents and reservoir equilibration processes for Kilauea magma. Reservoir-equilibrated magma is ultimately erupted in the summit region or injected into the rifts, from which it may subsequently be discharged during subaerial and submarine eruptions.

Three early CO₂ samples from Sulphur Bank fumarole just north-east of the caldera show a range of $\delta^{13}\text{C}$ values of –3.4 to –3.2‰ relative to the PDB standard (Table 1). D.M.T.² collected gas samples between 1975 and 1977 from Sulphur Bank and from the 1971 and 1974 fissure fumaroles formed during summit eruptions on the southern edge of the caldera, directly above the summit chamber. The temperature of the 1971 fissure fumarole was 140 °C and it emitted H₂O, CO₂ and SO₂ with small amounts of H₂S, N₂ and trace rare gases. The fumaroles sampled at Sulphur Bank and the 1974 fissure were at the local boiling point (96 °C) and emitted gases much poorer in sulphur. Sampling, extraction and analytical procedures have

Table 1 $\delta^{13}\text{C}$ values for Kilauea summit fumaroles

Site	Ref.	Date	$\delta^{13}\text{C}$ (‰)
Sulphur Bank	1	Not given	–3.2, –3.3, –3.4
Sulphur Bank	2	5 Aug. 1975	–3.19, –2.62
		18 Aug. 1975	–3.43
		8 Nov. 1975	–2.29, –3.24
		Dec. 1975	–3.01, –2.54
		13 Aug. 1976	–3.42
		17 Aug. 1976	–3.37
		13 Oct. 1976	–3.61, –3.01
		1 Dec. 1976	–2.68
		28 Dec. 1976	–3.51
		28 Jan. 1977	–3.3
1971 fissure	2	13 Aug. 1975	–3.35
		19 Aug. 1975	–3.21
		8 Nov. 1975	–3.29
		1 Dec. 1975	–3.19, –3.31
		9 Aug. 1976	–3.18, –3.28
		26 Aug. 1976	–3.41
		13 Oct. 1976	–2.81, –3.50
		1 Dec. 1976	–3.19
		28 Dec. 1976	–4.16
		29 Dec. 1976	–3.57
		29 Jan. 1977	–3.4
1974 fissure	2	29 Jan. 1977	–3.13
		10 Feb. 1977	–2.95

Table 2 Estimated $\delta^{13}\text{C}$ values for Kilauea chamber gas and parental magma

$\Delta^{13}\text{C}_{\text{g-m}}$	$\delta^{13}\text{C}_{\text{cg}}$	$\delta^{13}\text{C}_{\text{pm}}^*$	$\delta^{13}\text{C}_{\text{pm}}^\dagger$
3.8	–3.2	–3.5	–3.6
4.0	–3.0	–3.3	–3.4
4.3	–2.7	–3.0	–3.1

All units are ‰.

* (wt% CO₂)_{rem} = 0.049, (wt% CO₂)_{pm} = 0.65 (ref. 7).

† (wt% CO₂)_{rem} = 0.03, (wt% CO₂)_{pm} = 0.32 (ref. 8).

been described in detail elsewhere². The results (Table 1) yield an average $\delta^{13}\text{C}$ of –3.2‰. Duplicate samples showing $\delta^{13}\text{C}$ differences as large as those between samples collected several months apart are apparently caused by Rayleigh distillation processes related to cyclic accumulation and flashing of water at shallow depths².

Because most parental CO₂ is vented from the summit chamber^{7,8}, the $\delta^{13}\text{C}$ of parental magma may be similar to the fumarole value. We have tested this hypothesis with simple equilibrium and mass-balance models based on the assumptions: (1) that Rayleigh fractionation by continual loss and isolation of carbon in exsolved CO₂ is minimal during transport from the mantle and (2) that carbon in the chamber gas approaches isotopic equilibrium with carbon in co-genetic reservoir-equilibrated melt. Recent work³ has shown that Rayleigh fractionation effects on carbon isotopes are minimal during exsolution associated with the ascent and eruption of mid-oceanic basalt at seafloor depths >500 m. Thus, it is reasonable to assume that the same is true for the exsolution of CO₂ at considerably higher pressures from ascending Kilauea parental melt. Assumption (2) is supported by observations³ that the distribution of carbon isotopes between CO₂ in vesicles and the glass of mid-oceanic basalts corresponds closely to the experimentally determined equilibrium fractionation²².

For equilibrium fractionation of carbon isotopes between chamber gas and reservoir-equilibrated melt, the $\delta^{13}\text{C}$ for chamber gas ($\delta^{13}\text{C}_{\text{cg}}$) can be expressed as

$$\delta^{13}\text{C}_{\text{cg}} = \delta^{13}\text{C}_{\text{rem}} + \Delta^{13}\text{C}_{\text{g-m}} \quad (1)$$

where $\delta^{13}\text{C}_{\text{rem}}$ is the $\delta^{13}\text{C}$ for reservoir-equilibrated melt and $\Delta^{13}\text{C}_{\text{g-m}}$ is an approximation to the ‰ fractionation factor (1,000 ln $\alpha_{\text{g-m}}$) for carbon isotopes between CO₂ gas and basaltic melt. The value for $\Delta^{13}\text{C}_{\text{g-m}}$ has been determined from experimental products²² and samples of seafloor basalt³. The experimental data²² give a mean value of 4.3 for seven experiments over a temperature range of 1,120–1,280 °C at 7–8.5 kbar; temperature dependence is insignificant at the 95% confidence level. The mean $\Delta^{13}\text{C}_{\text{g-m}}$ calculated from isotope data for four seafloor basalts³ is 3.8. Therefore, we have used $\Delta^{13}\text{C}_{\text{g-m}} = 3.8\text{--}4.3$ in our calculations. Recent carbon isotope data for submarine lavas along the east rift³ provide an estimate for $\delta^{13}\text{C}_{\text{rem}}$. The $\delta^{13}\text{C}$ values for these flows at depths >500 m vary from –9 to –5‰, with a mean value of –6.2‰. The data support the interpretation³, however, that most of the lavas are derived from shallower portions of the rift storage system, which tend to accumulate CO₂-rich bubbles that are either carried into the rift from the summit chamber or exsolved during subsequent crystallization or depressurization associated with transport in the rift. As a result of ¹³C-enrichment in CO₂ gas relative to melt, the mean $\delta^{13}\text{C}$ for these flows is thought to be too high to be characteristic of the bulk of the reservoir-equilibrated magma³. Therefore, in our calculations we have used the intermediate value of –7‰ for $\delta^{13}\text{C}_{\text{rem}}$.

Our results (Table 2) are similar to the $\delta^{13}\text{C}$ data for the summit fumaroles (Table 1), suggesting that CO₂ degassed from the summit chamber is the principal source of summit fumarole

carbon. Moreover, they suggest that carbon isotopes in CO₂ vented from the summit chamber experience little fractionation in transit to the summit fumaroles, despite evidence for some carbonate alteration product formation in the hydrothermal system overlying the summit chamber²³.

The $\delta^{13}\text{C}$ for the parental magma ($\delta^{13}\text{C}_{\text{pm}}$) can be estimated from the conservation of CO₂ and carbon isotopes among parental magma, reservoir-equilibrated magma and chamber gas. This approach yields the mass-balance equation

$$\delta^{13}\text{C}_{\text{pm}} = \delta^{13}\text{C}_{\text{cg}} + \frac{(\text{wt}\% \text{CO}_2)_{\text{rem}}}{(\text{wt}\% \text{CO}_2)_{\text{pm}}} (\delta^{13}\text{C}_{\text{rem}} - \delta^{13}\text{C}_{\text{cg}}) \quad (2)$$

We have evaluated equation (2) for the same set of data as was used to calculate $\delta^{13}\text{C}_{\text{cg}}$. The CO₂ concentrations are based on two models for the volatile contents and degassing of Kilauea magma^{7,8}: one model⁷ gives 0.65 wt% CO₂ for the parental melt and 0.049 wt% CO₂ for reservoir melt equilibrated at an average depth of 3 km; the other⁸ gives corresponding values of 0.32 wt% CO₂ and 0.03 wt% CO₂. Because the concentration ratio in equation (2) is similar for each model, the estimates for $\delta^{13}\text{C}_{\text{pm}}$ summarized in Table 2 are in close agreement and lie in the range -3.6 to -3.0‰.

The $\delta^{13}\text{C}_{\text{pm}}$ values clearly support the hypothesis that carbon in parental magma isotopically resembles fumarole carbon and is similarly enriched in ^{13}C compared with mid-oceanic basalt ($\delta^{13}\text{C} = -4.5$ to -8‰)^{3,4,6} and with recent estimates for mantle carbon ($\delta^{13}\text{C} = -5$ to -7‰)³. Performing the calculations for $\delta^{13}\text{C}_{\text{rem}} = -6.2\text{‰}$, the mean value for east-rift submarine basalts³, gives even heavier parental carbon. If some Rayleigh fractionation is assumed during vapour exsolution, the estimated value for $\delta^{13}\text{C}_{\text{pm}}$ is also increased. Moreover, note that for $\delta^{13}\text{C}_{\text{pm}}$ to have a value in the range considered characteristic of mantle carbon, the parental CO₂ concentration would have to be reduced to ~0.07 wt%, which would imply summit CO₂ emissions of ~120 tons per day—that is, more than an order of magnitude lower than the observed rate⁸. Finally, fractionation and contamination mechanisms that may operate in the hydrothermal/fumarole system either are unable to transform carbon with $\delta^{13}\text{C}$ in the range considered characteristic of the mantle to that with $\delta^{13}\text{C}$ characteristic of the fumaroles, or they require removal of unrealistically large amounts of CO₂ as carbonate².

Turning to sulphur, recent work^{9,10} provides $\delta^{34}\text{S}$ data for east-rift submarine basalts, subaerial Kilauea lava (flows, spatter tephra, bombs), shallow dykes and summit fumarole gases (the 96 °C Jagger well fumarole at Sulphur Bank, a 128 °C 1971 fissure fumarole and a 145 °C 1974 fissure fumarole). $\delta^{34}\text{S}$ as used here refers to total sulphur (sulphide plus sulphate) in ‰ relative to the Canyon Diablo troilite standard. In reinterpreting these data, we assume Rayleigh fractionation during subaerial eruptive degassing, and estimate the fractionation factor ($\alpha_{\text{g-m}}$) for sulphur isotopes between gas and melt phases from the equation

$$\delta^{34}\text{S}_{\text{rem}} - \delta^{34}\text{S}_{\text{rd}} = 1,000 \ln f_{\text{g-m}}^{\alpha-1} \quad (3)$$

The weight fraction of sulphur remaining in lava after eruptive degassing ($f = 0.214$) is calculated from the residual sulphur content of glassy fountain spatter (150 parts per 10⁶, p.p.m.)²⁴ and the sulphur content of the reservoir-equilibrated melt (700 p.p.m.)^{7,9,10,25}. $\delta^{34}\text{S}_{\text{rem}}$ values fall in a narrow range around 0.7‰, based on data for east-rift submarine basalts^{9,10}. $\delta^{34}\text{S}_{\text{rd}}$, the analogous quantity for the residual sulphur, has a value of -0.4‰, based on the mean of data reported for spatter, bomb and shallow dyke samples⁹. The $\delta^{34}\text{S}_{\text{rd}}$ value used here does not include data for lava flows, which can continue to degas during flow after release of eruptive gases²⁴. The resulting value for $\alpha_{\text{g-m}}$ is 1.00075.

We used an equation analogous to equation (1), with $\Delta^{34}\text{S}_{\text{g-m}} = 1,000 \ln \alpha_{\text{g-m}} = 0.75$, to calculate $\delta^{34}\text{S}_{\text{cg}} = 1.5\text{‰}$. From an equation analogous to equation (2), we estimate that $\delta^{34}\text{S}_{\text{pm}} = 1.1\text{‰}$, in agreement with data for mid-oceanic basalts ($\delta^{34}\text{S} = -1$

to 2‰)^{11,26-28}. The $\delta^{34}\text{S}_{\text{cg}}$ value is higher than that measured for the 1971 and 1974 fissure fumaroles (0.9‰)⁹ but considerably lower than that observed for Sulphur Bank (8.0‰)⁹. The high value at the Sulphur Bank boiling-point fumarole is probably the result of fractionation associated with precipitation of sulphur by reaction of H₂S and SO₂ in the presence of liquid water. We attribute the difference in $\delta^{34}\text{S}$ values between the chamber gas and fissure fumaroles to loss of heavy sulphur during sulphate formation in hydrothermal processes adjacent to the summit chamber. Sulphate formation by disproportionation of chamber-gas SO₂ would produce fumarole gases relatively depleted in ^{34}S compared with the chamber gas²⁹. Loss of chamber-gas SO₂ in this manner is consistent with the observations that measured abundances of SO₂ emitted in the summit region^{30,31} are smaller than SO₂ discharges predicted for the summit chamber⁷.

Thus, we conclude that Kilauea parental magma is significantly enriched in ^{13}C compared with mid-oceanic basalt and in contrast to previous estimates of the isotopic composition of mantle carbon, although it is similar to mid-oceanic basalt in sulphur isotope composition. Moreover, massive chamber degassing of CO₂ effectively buffers the isotopic composition of carbon in the summit hydrothermal system and the summit fumaroles at values close to that of the parental magma, while sulphur isotopes in the chamber-gas discharge are significantly fractionated by reactions in the hydrothermal system.

The simplest explanation for enrichment of heavy carbon in Kilauea parental magma is a mantle with isotopically heterogeneous carbon, thereby permitting sources of hotspot and mid-oceanic basalt to differ in $\delta^{13}\text{C}$. Most models for the petrogenesis of Hawaiian tholeiite infer volatile components to have a deep, undepleted mantle source^{16,32}. The present results indicate that the source is enriched in heavy carbon, which in turn suggests an oxidized chemical form such as carbonate or CO₂. To render our results consistent with a mantle in which carbon is isotopically homogeneous⁶, it is necessary to assume (1) that carbon in the mantle and in the parental magmas of mid-oceanic basalt is heavier than the carbon observed in the seafloor lavas and (2) that an early stage of degassing analogous to Kilauea summit-chamber degassing removes heavier carbon from mid-oceanic basalt before it is extruded on the sea floor.

This work was supported at Sandia National Laboratories by the Office of Basic Energy Research, US Department of Energy under contract DE-AC04-76DP00789. We thank H. Craig, L. P. Greenland and T. J. Casadevall for stimulating discussions.

Received 19 July; accepted 31 October 1985.

1. Craig, H. in *Nuclear Geology on Geothermal Areas* (ed. Tongiorgi, E.) 17-53 (Consiglio Nazionale Della Ricerche, Laboratorio Di Geologia Nucleare, Pisa, 1963).
2. Thomas, D. M. thesis, Univ. Hawaii (1977).
3. Des Marais, D. J. & Moore, J. G. *Earth planet. Sci. Lett.* **69**, 43-57 (1984).
4. Pineau, F., Javoy, M. & Bottinga, Y. *Earth planet. Sci. Lett.* **29**, 413-421 (1976).
5. Deines, P. *Geochim. cosmochim. Acta* **44**, 943-961 (1980).
6. Pineau, F. & Javoy, M. *Earth planet. Sci. Lett.* **62**, 239-257 (1983).
7. Gerlach, T. M. & Graeber, E. J. *Nature* **313**, 273-277 (1985).
8. Greenland, L. P., Rose, W. I. & Stokes, J. B. *Geochim. cosmochim. Acta* **49**, 125-129 (1985).
9. Sakai, H., Casadevall, T. J. & Moore, J. G. *Geochim. cosmochim. Acta* **46**, 729-738 (1982).
10. Sakai, H., Des Marais, D. J., Ueda, A. & Moore, J. G. *Geochim. cosmochim. Acta* **48**, 2433-2441 (1984).
11. Puchelt, H. & Hubberten, H.-W. *Init. Rep. DSDP* **51**, **52**, **53**, **54**, Pt 2, 1145-1148 (1980).
12. Swanson, D. A., Duffield, W. J. & Fiske, R. S. *U.S. geol. Surv. Prof. Pap.* **963**, 1-39 (1976).
13. Ryan, M. P., Blevins, J. Y. K., Okamura, A. T. & Koyanagi, R. Y. *J. geophys. Res.* **88**, 4147-4181 (1983).
14. Eaton, J. P. in *The Crust of the Pacific Basin* (ed. Endo, E. T.) 13-29 (American Geophysical Union, 1962).
15. Ellsworth, W. L. & Koyanagi, R. Y. *J. geophys. Res.* **82**, 5379-5394 (1977).
16. Wright, T. L. *J. geophys. Res.* **89**, 3233-3252 (1984).
17. Ryan, M. P., Koyanagi, R. Y. & Fiske, R. S. *J. geophys. Res.* **86**, 7111-7129 (1981).
18. Wright, T. L. & Fiske, R. S. *J. Petrol.* **12**, 1-65 (1971).
19. Duffield, W. A., Christiansen, R. L., Koyanagi, R. Y. & Peterson, D. W. *J. Volcanol. geotherm. Res.* **13**, 273-307 (1982).
20. Dzurison, D., Koyanagi, R. Y. & English, T. T. *J. Volcanol. geotherm. Res.* **21**, 177-206 (1984).
21. Macdonald, G. A. & Abbott, A. T. *Volcanos in the Sea* (University Press of Hawaii, 1974).
22. Javoy, M., Pineau, F. & Iiyama, L. *Contr. Miner. Petrol.* **67**, 35-39 (1978).
23. Keller, G. V., Grose, L. T., Murray, J. C. & Skokan, C. K. *J. Volcanol. geotherm. Res.* **5**, 345-385 (1979).
24. Swanson, D. A. & Fabbri, B. P. *J. Res. U.S. Geol. Surv.* **1**, 649-658 (1973).
25. Moore, J. G. & Fabbri, B. P. *Contr. Miner. Petrol.* **33**, 118-127 (1971).
26. Kanehira, K., Yui, S., Sakai, H. & Sasaki, A. *Geochemistry* **7**, 89-96 (1973).

27. Grinenko, V. A., Dmitriev, L. V., Migdisov, A. A. & Sharaskin, A. Ya. *Geochem. Int.* **1975**, 132–137 (1975).
28. Sakai, H., Ueda, A. & Field, C. W. in *Short Pap. 4th int. Conf. Geochronology, Cosmochemistry and Isotope Geology* (ed. Zartman, R. E.) 372–374 (US Geological Survey Open-File Report 78–701, 1978).
29. Ohmoto, H. & Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 509–567 (Wiley, New York, 1979).
30. Stoiber, R. E. & Malone, G. B. *EOS* **56**, 461 (1975).
31. Casadevall, T. J. & Hazlett, R. W. *J. Volcanol. geotherm. Res.* **16**, 173–188 (1983).
32. Sen, G. *Earth planet. Sci. Lett.* **62**, 215–228 (1983).

Calcium anomalies in the mantle and a subducted metaserpentinite origin for diamonds

Daniel J. Schulze*

Geophysical Laboratory, Carnegie Institution of Washington, Washington, DC 20008, USA

Most natural diamonds are thought to have formed within the upper mantle, before transport to the surface by kimberlites or lamproites, in a magnesian, calcium-poor, garnet-bearing harzburgite or dunite. Serpentinites, formed by hydration of oceanic lithosphere, are sufficiently poor in calcium and have high enough Mg/(Mg + Fe) values that, following subduction and prograde metamorphism to ~1, 100 °C, 55–60 kbar, they would equilibrate as such harzburgites and dunites. Diamonds may be the high-pressure equivalent of graphite produced, or introduced, during serpentinization. The apparent restriction of diamonds and low-Ca garnet xenocrysts to cratons may result from the fact that subducted metaserpentinites, lodged within the mantle beneath the cratons, have been effectively isolated from participation in subsequent tectonic and magmatic events that occurred outside the craton margins.

The most important constraint on the origin of diamonds is imposed by the composition of the minute syngenetic mineral inclusions that occur in some diamonds. These minerals are considered to have been in stable coexistence with their host diamond at the time of their inclusion^{1,2}, so their compositions can yield valuable information concerning the chemical environment and ambient pressure and temperature during diamond formation. The composition of these mineral inclusions has been well documented (see refs 1–4). Most syngenetic mineral inclusions belong to one of two suites, a peridotitic suite and an eclogitic suite, corresponding to the two most abundant types of ultramafic xenoliths found in kimberlites. Just as peridotite xenoliths are more abundant than eclogite xenoliths, the peridotitic inclusions in diamonds are more common than are eclogitic suite inclusions.

There are, however, distinct and important differences in paragenesis and chemical composition between the minerals included in diamonds and most of their xenolithic counterparts. Within the peridotite suite of diamond inclusion minerals diopside is rare. Most assemblages represent a garnet-bearing dunite or harzburgite paragenesis (olivine ± garnet ± orthopyroxene ± chromite), which is also reflected in the Mg-rich and Ca-poor mineral compositions. Most peridotitic garnets and orthopyroxenes found as inclusions in diamonds have calcium contents too low to have been in equilibrium with diopside⁵. This situation is in stark contrast to the fact that garnet lherzolites are the most abundant variety of upper mantle xenolith found in kimberlites⁶. Furthermore, peridotite minerals found as inclusions in diamonds are generally more magnesian than their counterparts in the common garnet peridotite xenoliths. For example, olivine inclusions in diamonds generally have

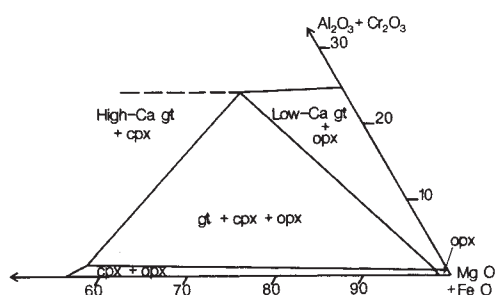


Fig. 1 Phase compatibility diagram (modified ACF plot) for ultrabasic compositions equilibrated near 1,100 °C, 55–60 kbar (mol %). Data used to construct this diagram were taken from the compositions of minerals associated with diamonds from the Finsch Mine in South Africa^{2,4,13,44} and the experimental data of Lindsley and Dixon⁴⁵.

Mg/(Mg + Fe) values of 0.92–0.95 (refs 2, 7), whereas olivines in garnet peridotites most often have values of 0.90–0.93 (see refs 8–10).

Diamond-bearing peridotites also occur, although only about 20 are known. Four of these are garnet lherzolite that are virtually indistinguishable in texture, mode and mineral chemistry from the abundant diamond-free garnet lherzolite xenoliths from kimberlites^{11–13}. Most diamondiferous peridotites, however, are garnet- and chromite-bearing harzburgites, and dunites, the minerals of which have the same chemical characteristics as those that occur as inclusions in diamonds^{12,14,15}.

Thus, diamonds with peridotite affinity seem to occur almost exclusively in a magnesian, Ca-deficient peridotite source rock, a garnet harzburgite or garnet dunite that is chemically unlike most of the upper mantle. Two of the outstanding problems regarding the genesis of diamond are therefore: (1) how a calcium-deficient peridotite acquires such unusual chemical characteristics in a 'diopside-saturated' upper mantle; and (2) why diamonds are found almost exclusively in this rock type, and generally not in the more abundant garnet lherzolites.

Students of ultramafic rocks occurring in the Earth's crust are familiar with rocks of both anomalously low and high calcium contents. During serpentinization of peridotite, a process restricted to lower temperature crustal conditions, calcium is removed from pyroxenes and is deposited in associated non-peridotite lithologies—dykes, country rock or tectonic inclusions¹⁶. The rocks resulting from this process are serpentinite, with a very low calcium content, and related rodingite, with an anomalously high calcium content. It is proposed that subduction of serpentinites, formed by hydration of oceanic lithosphere, and subsequent metamorphism to garnet peridotite facies may be responsible for upper mantle garnet peridotites that are depleted in calcium relative to most of the upper mantle. Several other workers have previously suggested that certain upper mantle xenoliths in kimberlites formed from subducted oceanic crust^{17–22}. These models are compared with the present one below.

To test the hypothesis that serpentinite could be the chemical equivalent of Ca-poor, garnet-bearing harzburgite in the upper mantle, a phase compatibility diagram has been constructed (Fig. 1), corresponding to natural peridotite at ~1,100 °C and 55–60 kbar, the approximate equilibration conditions of diamond inclusions and diamond-bearing peridotites from the Finsch kimberlite in South Africa¹³. By plotting serpentinite bulk chemical compositions, the serpentinites are effectively recast into approximate proportions of the various minerals that are stable in these conditions.

Most of the natural compositions used to construct Fig. 1 were taken from the data of Gurney and co-workers for the Finsch pipe in South Africa^{2,4,13,44,45}. These were used because they represent the best documentation of diamond-related

* Present address: Department of Geological Sciences, Queen's University, Kingston, Ontario, Canada K7L 3N6.

minerals from a single pipe, and because two diamond-bearing garnet lherzolites have been described from this pipe¹³.

Projected onto the CaO- and (Al₂O₃ + Cr₂O₃)-poor area of this diagram, expanded in Fig. 2, are the bulk compositions of 57 serpentinites taken from the literature. Most are from Alpine-type serpentinites and abyssal serpentinites and are massive rocks, although some vein serpentinites have been included. Figure 2 shows that most (42) of these serpentinites are compositionally equivalent to harzburgites in which a low-calcium garnet is stable. The remainder of the compositions plot in the fields of garnet lherzolite and garnet-free harzburgite.

The Mg/(Mg + Fe) values of the 42 serpentinites that correspond to low-calcium, garnet-bearing peridotites on this diagram range from 0.88 to 0.98, although most of them have values between 0.92 and 0.95, with a mean value of 0.931. This value is well within the range of Mg/(Mg + Fe) values of olivines that occur as inclusions in diamonds.

On the basis of the similarities in chemical composition between the serpentinites and the minerals included in diamonds, it is concluded that serpentinite has the requisite bulk composition to be a suitable host rock for diamonds of peridotite affinity, if subducted under the continent and equilibrated at ~1,100 °C and 55–60 kbar. These pressure and temperature conditions correspond roughly to 180–200 km beneath a stable shield, on a continental geothermal gradient¹³.

Several of the serpentinite analyses plot in the garnet lherzolite field of Fig. 2. Their relatively high calcium contents are probably due to the presence of relict diopside that was not completely replaced by serpentine. Such rocks could be protoliths of diamond-bearing garnet lherzolites.

Theories on the origin of diamonds, centred around the need to explain the calcium-poor nature of the included peridotite minerals, are (1) that diamonds are phenocrysts in an early stage of the kimberlite magma, and the Ca-poor nature of their associated minerals is due to complexing of the calcium in the diopside of a garnet lherzolite parent by partial melting in the presence of CO₂ (ref. 34), and (2) that the low-Ca garnet harzburgite is a residue remaining after extraction of komatiite liquids from parental garnet lherzolite in Archaean time³⁵. Richardson *et al.*³⁶ have essentially combined both of these theories, although rejecting the hypothesis that the diamonds are actual phenocrysts in the same kimberlite that carried them to the surface, because of the age difference between the young kimberlites (~90 Myr, ref. 37) and the Archaean Sm–Nd model ages of garnet inclusions in diamonds³⁶.

The komatiite residue models are based on the apparent restriction of komatiites, diamonds and low-Ca garnets to stable Archaean cratons, and on the assumption that a garnet harzburgite or garnet dunite would remain as a residue of partial melting of peridotite to yield komatiite. Phase-equilibrium studies, however, have shown that during high degrees of partial melting of garnet peridotite, either garnet or diopside may be the first phase to be consumed, depending on the original bulk composition³⁸. The spatial association of low-Ca garnets and komatiites³⁵ may, therefore, be coincidental.

In the serpentinite model for the origin of peridotite-suite diamonds, the apparent restriction of Ca-poor garnet harzburgites to stable cratons (such as Kaapvaal Craton in South Africa³⁵) may be due to the fact that following serpentinite emplacement, perhaps through shallow subduction^{39,40}, the sub-cratonic mantle remained 'cool'. The upper mantle in adjacent mobile belts, however, has been remobilized and heated by tectonism and magma production during subsequent orogenesis. During such events, diamond-bearing metaserpentinites may lose both their diamonds and their low-Ca 'signature' through heating and hybridization (for example, ref. 41).

An origin involving a subducted protolith has previously been invoked for several types of xenoliths, including diamonds^{17–22}. The model proposed here is similar in several respects to two of these earlier models, which are discussed below.

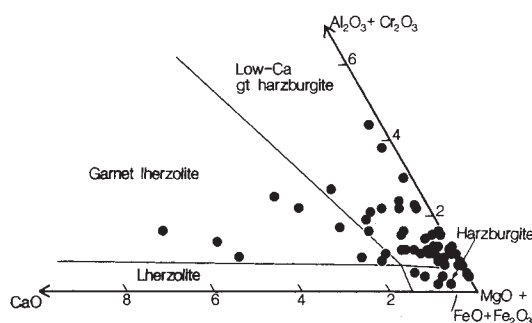


Fig. 2 Bulk serpentinite compositions, plotted in an expanded portion of the lower right-hand side of Fig. 1. Serpentinites taken from refs 16, 23–33.

Ringwood¹⁹ proposed that diamonds originate in subducted oceanic spinel-bearing harzburgite cumulates. Although serpentinitization is included in his model, its role is minor, and apparently was included only to facilitate partial melting during subduction, yielding the residual character of the subducted peridotite. The low-calcium nature of the harzburgite in Ringwood's model is attained through igneous processes, with no major role attached to the serpentinitization. In the present model, the precise chemical nature of the peridotite protolith, especially its calcium content, before serpentinitization, is unimportant. Thorough serpentinitization removes Ca from the bulk rock so efficiently that even a lherzolite parent can yield a Ca-poor serpentinite product (see ref. 23). A cumulate harzburgite as proposed by Ringwood¹⁹, therefore, is unnecessary.

Helmstaedt and co-workers^{20,21} have proposed that grosspyrite xenoliths (grossular garnet, omphacitic clinopyroxene and kyanite) are the prograde products of metamorphosed subducted rodingites. This process is the complement of that proposed here. They also suggested that the serpentinites subducted together with rodingites would form garnet peridotites in the mantle, although they did not explore the possibility that serpentinites could be the chemical equivalents of the low-calcium garnet peridotites.

Although it is generally assumed that carbon is present in the mantle as graphite or diamond, graphite-bearing peridotites are even less abundant than are the very rare diamond-bearing samples (F. R. Boyd, personal communication). It appears, therefore, that the source of the carbon is linked with the processes that deplete peridotite in calcium, which in this model is serpentinitization.

Pasteris⁴² has shown that graphite can occur in serpentinitized olivine. Whether it is formed from CO₂ inclusions in the pre-existing olivine or from methane-bearing fluids in the reducing conditions that exist during serpentinitization⁴³, or was present in the fresh olivine itself⁴⁴ and was only released as graphite during serpentinitization, is not clear. It is possible, however, that the carbon that eventually formed as diamond originated in the process of serpentinitization, as did the chemical signature of the calcium-deficient peridotite host rock.

I thank F. R. Boyd, H. Helmstaedt, B. O. Mysen and H. S. Yoder Jr, for helpful reviews and M. Imlay for technical assistance.

Received 19 June; accepted 15 October 1985.

1. Meyer, H. O. A. & Boyd, F. R. *Geochim. cosmochim. Acta* **36**, 1255–1273 (1972).
2. Gurney, J. J., Harris, J. W. & Rickard, R. S. in *Kimberlites, Diamonds and Diatremes: Their Geology, Petrology and Geochemistry* (eds Boyd, F. R. & Meyer, H. O. A.) 1–15 (American Geophysical Union, Washington DC, 1979).
3. Prinz, M., Manson, D. V., Hlava, P. F. & Keil, K. *Phys. Chem. Earth* **9**, 797–815 (1975).
4. Tsai, H. M., Meyer, H. O. A., Moreau, J. & Milledge, H. J. in *Kimberlites, Diamonds and Diatremes: Their Geology, Petrology and Geochemistry* (eds Boyd, F. R. & Meyer, H. O. A.) 16–26 (American Geophysical Union, Washington, DC, 1979).
5. Sobolev, N. V., Lavrent'ev, Y. G., Pokhilenko, N. P. & Usova, L. V. *Contr. Miner. Petrol.* **40**, 39–52 (1973).
6. Mathias, M., Siebert, J. C. & Rickwood, P. C. *Contr. Miner. Petrol.* **26**, 75–123 (1970).
7. Meyer, H. O. A. & Tsai, H. M. *Miner. Sci. Engng* **8**, 242–261 (1976).
8. Nixon, P. H. & Boyd, F. R. in *Lesotho Kimberlites* (ed. Nixon, P. H.) 48–56 (Cape and Transvaal, Capetown, 1973).

9. Boyd, F. R. & Nixon, P. H. *Phys. Chem. Earth* **9**, 431-454 (1975).
10. Danchin, R. V. in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics* (eds Boyd, F. R. & Meyer, H. O. A.) 104-126 (American Geophysical Union, Washington, DC, 1979).
11. Dawson, J. B. & Smith, J. V. *Nature* **254**, 580-581 (1975).
12. Pokhilenko, N. P., Sobolev, N. V. & Lavrent'ev, Yu. G. in *Extended Abstr. 2nd int. Kimberlite Conf.*, Santa Fe (American Geophysical Union, Washington, DC, 1977).
13. Shee, S. R., Gurney, J. J. & Robinson, D. N. *Contr. Miner. Petrol.* **81**, 79-87 (1983).
14. McCallum, M. E. & Egger, D. H. *Science* **192**, 253-256 (1976).
15. Boyd, F. R. & Finnerty, A. A. *J. geophys. Res.* **85**, 6911-6918 (1980).
16. Coleman, R. G. *Bull. N.Z. geol. Surv.* **76**, 102 pp.
17. Sharp, W. E. *Earth planet. Sci. Lett.* **21**, 351-354 (1974).
18. Helmstaedt, H. & Doig, R. *Phys. Chem. Earth* **9**, 95-111 (1975).
19. Ringwood, A. E. *Earth planet. Sci. Lett.* **36**, 443-448 (1977).
20. Helmstaedt, H. *EOS* **59**, 392-393 (1978).
21. Helmstaedt, H., Carmichael, D. M. & Percival, J. A. *Abstr. Prog. geol. Soc. Am.* **11**, 442 (1979).
22. Exley, R. A., Smith, J. V. & Dawson, J. B. *Am. Miner.* **68**, 512-516 (1983).
23. Nicolas, A. *Bull. volcan.* **32**, 499-508 (1969).
24. Aumento, F. & Loubat, H. *Can. J. Earth Sci.* **8**, 631-663 (1971).
25. Chidester, A. H. *U.S. geol. Surv. Prof. Pap.* **345** (1962).
26. Coleman, R. G. & Keith, T. E. *J. Petrol.* **12**, 311-328 (1971).
27. Francis, G. H. *Am. J. Sci.* **254**, 201-226 (1956).
28. Grubb, P. L. C. *Econ. Geol.* **57**, 1228-1246 (1962).
29. Hess, H. H., Smith, R. J. & Dengo, G. *Am. Miner.* **37**, 68-75 (1952).
30. Larrabee, D. M. *Bull. U.S. geol. Surv.* **1283** (1969).
31. Loferski, P. J. thesis. Virginia Polytechnic Inst. (1980).
32. Mumpton, F. A. & Thompson, C. S. *Clays Clay Miner.* **23**, 131-143 (1975).
33. Prinz, M. *et al. J. geophys. Res.* **81**, 4087-4103 (1976).
34. Harte, B., Gurney, J. J. & Harris, J. W. *Contr. Miner. Petrol.* **72**, 181-190 (1980).
35. Boyd, F. R. & Gurney, J. J. *Yb. Carnegie Instn Wash.* **81**, 261-267 (1982).
36. Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198-202 (1984).
37. Davis, G. L. in *Extended Abstr. 2nd int. Kimberlite Conf.*, Santa Fe (American Geophysical Union, Washington DC, 1977).
38. Davis, B. T. C. & Schairer, J. F. *Yb. Carnegie Instn Wash.* **64**, 123-126 (1965).
39. Vajner, V. J. A. *Rep. 1972-73, Precamb. Res. Unit, Univ. Cape Town*, 12-23 (University of Cape Town, 1974).
40. Helmstaedt, H. & Gurney, J. J. in *Kimberlites I: Kimberlites and Related Rocks* (ed. Kornprobst, J.) 425-434 (Elsevier, Amsterdam, 1984).
41. Wyllie, P. J. & Sekine, T. *Contr. Miner. Petrol.* **79**, 375-380 (1982).
42. Pasteris, J. D. *Geology* **9**, 356-359 (1981).
43. Freund, F., Kathrein, H., Wengeler, H., Knobel, R. & Heinen, H. *J. Geochim. cosmochim. Acta* **44**, 1319-1333 (1980).
44. Gurney, J. J. & Switzer, G. S. *Contr. Miner. Petrol.* **39**, 103-116 (1973).
45. Lindsley, D. H. & Dixon, S. A. *Am. J. Sci.* **276**, 1285-1301 (1976).

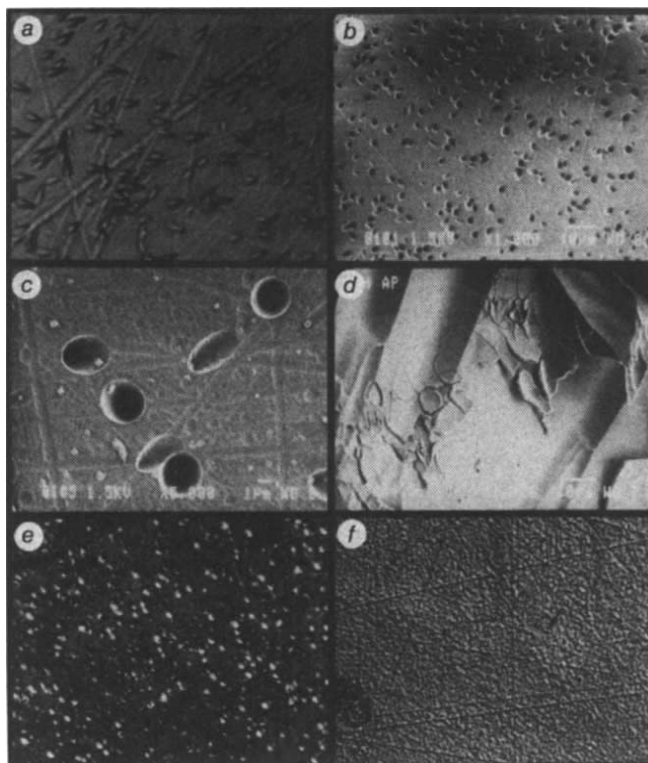


Fig. 1 *a-d*, Fission tracks in simple silicate glass 1 (SiO_2 , 54.5%; B_2O_3 , 29.2%; Na_2O , 10.3%; Li_2O , 4.9%), etched in deionized water. *a*, Transmission optical micrograph, 15 h at 50 °C (average track length 10 μm); *b*, *c*, SEM micrograph, 60 min at 100 °C; *d*, SEM micrograph of a grain leached for 5 days at 50 °C clearly showing the formation of the hydrated layer. Note the sharp smooth interface with the pristine glass, its fissuring as well as its peeling off. Completely similar results were also obtained with glass 2 (SiO_2 , 64%; B_2O_3 , 12%; Na_2O , 24%). H- and Na-depth profiling by means of specific resonant nuclear reactions, respectively $^1\text{H}(^{15}\text{N}, \alpha\gamma)^{12}\text{C}$ and $^{23}\text{Na}(p, \alpha\gamma)^{20}\text{Ne}$, allowed us¹⁵ to demonstrate ion exchanges in sodium-containing glasses leached at low temperatures (<100 °C) for short periods of time and developing a thin (a few thousand Å) hydrated layer. Direct visual observations for samples leached at a higher temperature (160 °C) show that a 1-cm-thick layer can easily be formed¹⁷. *e*, *f*, Fission tracks in complex glass 3 (SiO_2 , 54%; B_2O_3 , 13.5%; Na_2O , 13%; CaO , 5%; various transition and heavy-element oxides, 14%) and complex glass 4 (SiO_2 , 46.1%; Al_2O_3 , 5%; B_2O_3 , 14.2%; Na_2O , 12.5%; CaO , 4.1%; Li_2O , 2%; P_2O_5 , 0.24%; various oxides, 15.9%) leached at 100 °C during 423 h in H_2O and 33 h in 250 g l^{-1} NaCl solution, respectively.

Methods. Polished sections of different borosilicate glasses were first irradiated with a ^{252}Cf source at total doses of 10^6 – 10^9 FF cm^{-2} , then placed in Teflon bottles and subjected to dissolution in deionized water, NaCl brine (250 g l^{-1}) or an HF solution (0.4 wt%) at temperatures of 20, 50 and 100 °C, and for times ranging from a few hours to several days. Observations were made on an optical microscope fitted with a Nomarski interference contrast device and on a Jeol JSM 840 scanning electron microscope, allowing a direct observation of samples at low accelerating voltage without any conducting surficial coating.

Mechanism of aqueous dissolution of silicate glasses yielded by fission tracks

J.-C. Dran*, J.-C. Petit† & C. Brousse†

* Laboratoire René Bernas, CNRS, BP 1, 91406 Orsay, France

† SESD/LECALT, CEA, CEN-FAR, BP 6, 92260 Fontenay-aux-Roses, France

Controversy continues over the mechanism of the aqueous corrosion of silicate glasses, largely because the processes by which the superficial hydrated layer is formed have not been clearly identified¹. Two alternative models are current. One view²⁻⁴ is that mobile ions in the glass network (such as Na^+ , Li^+ and K^+) are exchanged with hydrogen ions (such as H^+ and H_3O^+) and even H_2O in the solution⁵; the other is that the hydrated layer results from the re-precipitation from solution of dissolved network material⁶⁻¹². The first model is thought appropriate for simple man-made glasses and the second for more complex and natural glasses. We report here an optical and scanning electron microscope study of the corrosion of various borosilicate glasses irradiated by fission fragments. For all glass-solution combinations studied, we find that the hydrated layer embodies relics of the fission tracks, suggesting that this layer is mainly formed by *in situ* replacement, but in some cases a step re-precipitation is also apparent.

Our starting point was the assumption that it might be possible simply to discriminate between the two models by inducing structural microscale modifications of the glass surface and following their evolution during dissolution. We thought that these modifications could be introduced by external irradiation using low doses of fission fragments (FF) emitted by a ^{252}Cf source; these heavy ions are indeed able to produce in insulators very small linear zones of heavily damaged material known as latent tracks. If dissolution is selective, the hydrated layer is

mainly formed by an *in situ* transformation of the glass, in which case it may be possible to observe specific track remnants by choosing convenient experimental conditions. In contrast, in the case of congruent dissolution at the reaction interface, the hydrated layer is a re-precipitated coating that should gradually fill and cover the tracks. This simple idea requires that the local modifications have small dimensions compared with the thickness of the hydrated layer implied by dissolution in convenient experimental conditions; this is the case for latent fission tracks (length ~ 10 μm , diameter ~ 50 – 100 Å).

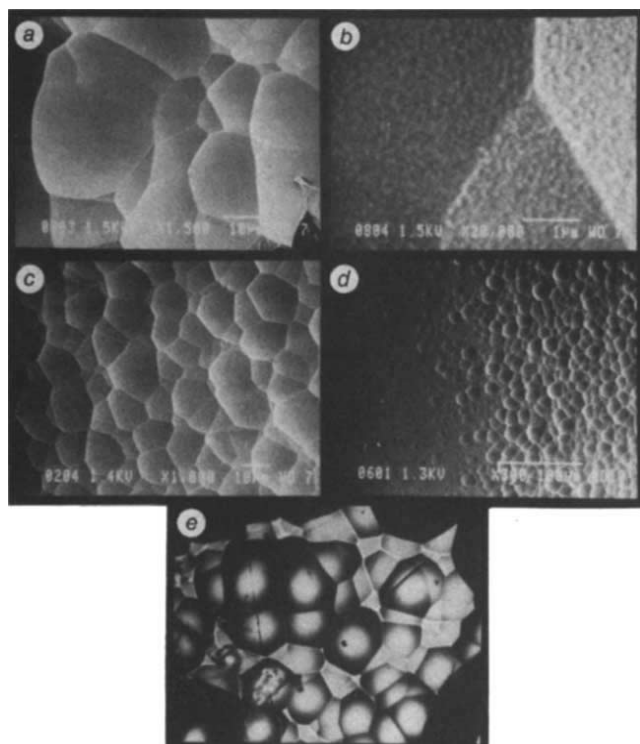


Fig. 2 *a, b*, SEM micrographs showing the cupule structure obtained in glass 1 after leaching for 5 days at 50 °C. *c, d*, A very similar surficial structure produced by prolonged dissolution of fission tracks in the same glass, 16 h at 100 °C (see text). Note in *d* the limit between irradiated (right) and unirradiated zones (left). *e*, Reflection optical micrograph of glass 3, leached for ~1 week at 100 °C in water at pH ~5.3, showing, after removal of the hydrated layer, a typical cupule structure.

We found that in simple glasses, etched fission tracks are clearly seen in most glass/solution combinations, and particularly in deionized water (Fig. 1). This result is interesting because, in such detectors, tracks are usually revealed only when very strong etchants, such as acid or alkaline solutions¹³, are used. Moreover, the scanning electron micrographs clearly show the presence of an apparently homogeneous hydrated layer ~1 µm thick; the fact that the ~10-µm-long tracks intersect this layer and keep perfect elliptical or circular openings argues strongly for the *in situ* formation of such a layer, and therefore for a selective dissolution. This conclusion agrees well with previous results obtained on this type of glass using various surface analytical techniques^{3,14,15}.

On continued etching, such that the thickness dissolved exceeds the range of the fission tracks, the etch pits develop a spherical shape, due to the isotropy of the etch rate in glasses, and begin to overlap. The resulting topography of the dissolution front (Fig. 2*c, d*), showing a typical cup-shaped ('cupule') structure, is very similar to that obtained during the normal dissolution of the same glass (Fig. 2*a, b*). The two hypotheses invoked to explain the formation of dissolution cupules are chemical/structural heterogeneities in the bulk of the glass and surficial defects, perhaps due to sample preparation techniques (sectioning and polishing). Clearly, the first hypothesis is the more interesting as, in this case, the cupule structure is related to the basic process of dissolution, which should therefore be congruent; the cupule should perpetuate itself as dissolution proceeds. In contrast, in the second hypothesis, cupules are merely artefacts, which should slowly disappear with sufficiently long dissolution times. Our results provide the first clear demonstration that surficial defects are indeed sufficient to propagate a structure of dissolution cupules at large depths.

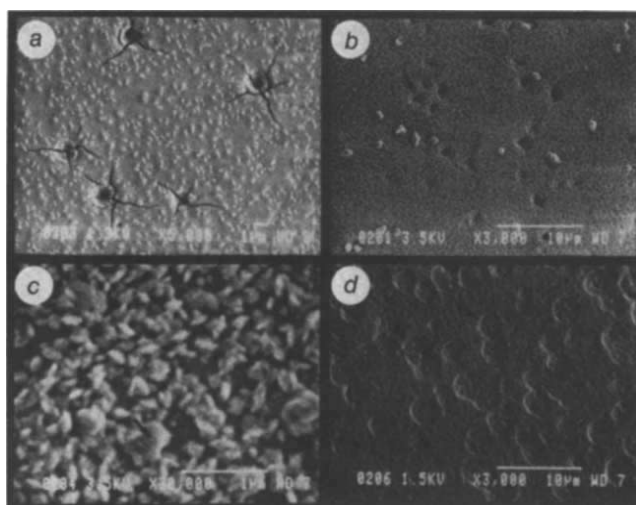


Fig. 3 Evolution during leaching in pure water of pre-etched fission tracks (5 min HF 1/100) in complex silicate glass 5 (SiO₂, 45.5%; Al₂O₃, 4.9%; B₂O₃, 14%; Na₂O, 9.9%; CaO, 4%; Fe₂O₃, 2.9%; ZnO, 2.5%; ZrO₂, 1%; Li₂O, 2%; P₂O₅, 0.28%; various transition and heavy-element oxides, 13%). *a*, SEM micrograph showing in the hydrated layer (leaching conditions, 69 h at 100 °C) remnants of FF tracks and their associated micro-cracks (see text); note the small granules on the surface. ×1,950. *b*, As in *a*, but after 240 h; the arrows indicate partially obstructed tracks and still open ones. Such a feature, combined with the absence of track-induced fracturing, argues strongly for a precipitated coating. ×1,170. *c*, Same field of view showing the granules at higher magnification. ×11,700. *d*, Track remnants in the inner hydrated layer after careful mechanical removal of the surficial sub-layer. ×1,170.

Therefore, cupules are probably artefacts produced by the surface state of the sample and are not necessarily linked to bulk microheterogeneities. An additional (and we think strong) argument is the absence of the 'cupule-within-cupule' structure which would occur when the dissolution front of one major cupule encounters new bulk heterogeneities. Moreover, although we agree that the spherical shape of the cupules implies a constant progression rate of the dissolution front, this is not necessarily indicative of a congruent dissolution. There are two arguments against this interpretation. First, high-resolution scanning electron micrographs (see Fig. 2*b*) show that the surface of the cupules is rough, rather like a hydrated silica gel, and not smooth as would be expected for a direct attack on the network. Second, both a constant progression rate and a sharp interface between the pristine glass and the hydrated layer can be simply explained by a marked increase of diffusivities of species in the layer relative to the glass (the hydrated layer is then almost 'transparent' for diffusion)^{16,17}. This seems likely if one takes into account the high porosities of such layers. In this hypothesis, the selective processes occurring at the interface are limited to very small thicknesses (for example <100–1,000 Å); this is compatible with the sharpness of the interface reported by various investigators^{6,8,9,11,12,18}.

In contrast, in complex glasses, the observations are more difficult to interpret because etched fission tracks are seen in some glass/solution combinations but not in others (see Fig. 1*e, f*), and there seems to be no simple correlation with glass/solution composition. We have already reported that such complexity^{19,20} exists in the behaviour of implanted surfaces with regard to aqueous corrosion and consider that this simply reflects our poor understanding of the dissolution processes in multi-component glasses. However, our results prove that fission tracks can, in many cases, be preferentially etched in complex glasses by solutions such as those found in geological environments. This is the first report of such an observation, and its implications

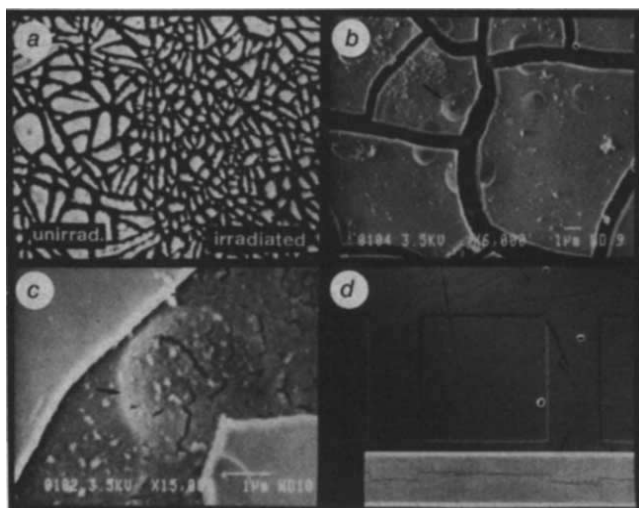


Fig. 4 Evolution during leaching in the NaCl brine of pre-etched fission tracks (5 min HF 1/100) in complex silicate glass 5. *a*, Reflection optical micrograph after 90 h at 100 °C (mean fracture width 1 μm); note the difference in fracture densities between irradiated and unirradiated areas (see text); the micro-cracking has often been reported in silica gel layers, but our observations show that for high doses of FF ($>10^7$ FF cm^{-2}), the density of this crack network, which mimics the image of the grid used as a mask during the irradiations, is markedly increased. As this phenomenon is probably an artefact of the observation technique, it may not affect the overall glass dissolution rate. *b*, *c*, SEM micrographs after 340 h at 100 °C showing track remnants in both sub-layers. This two-sub-layer structure has been described by others^{6,8,11,18}. In the case of glass 5, for instance, the inner sub-layer tends to accumulate elements such as Si, P, Zr, U, Th, rare-earth elements, Cr and Mn, whereas the external one preferentially concentrates elements forming very insoluble compounds, such as Ni; Fe and Zn are present in both sub-layers. In contrast, Na, B, Mo, Cs, Ca and Al are strongly depleted in the hydrated layer. In addition, in such complex glasses the hydrated layer contains tiny crystalline phases^{8,9,11}. *d*, The result of a complementary experiment in which a polished section of a glass had been implanted through an electron microscope grid with 207-keV lead ions (range ~ 500 Å) at a dose of 10^{13} ions cm^{-2} and then leached in the same NaCl solution. The example given is of glass 6 (SiO₂, 54%; Al₂O₃, 14%; B₂O₃, 9%; Na₂O + K₂O + MgO + Fe₂O₃, 2%; CaO, 21%), showing the formation of bumps ~ 200 Å high in the irradiated areas (width 100 μm), faithfully following the irradiation geometry (see corresponding step-height profile). This feature is explained by the preferential hydration and correlative swelling of the $\sim 1,000$ -Å-thick implanted material¹⁹ and not by a re-precipitated coating, which would not show the irradiation geometry so clearly.

for the resistance to aqueous corrosion of glasses containing fissionable elements (spontaneous or induced fission) should be carefully investigated.

In other experiments, we pre-etched the fission tracks in a dilute HF solution and further dissolved the samples in various mild solutions. Using the scanning electron microscope (SEM), we observe (see, for example, Figs 3 and 4 for glass 5) that the microstructure of the hydrated layer and therefore the evolution of the pre-etched holes strongly depend on the nature of the leachant. For example, in water, for leach times of ~ 69 h at 100 °C, the layer incorporates remnants of tracks, thereby confirming that it develops *in situ* and is not completely redeposited material. This conclusion is confirmed by the observation that stresses are induced inside the hydrated layer by each track and are seen as micro-cracks. These probably formed during the desiccation of the sample that was required for its observation in the vacuum chamber of the SEM. We

think that the tracks can induce stresses in the hydrated layer only if it develops around them; if the overall hydrated layer was a redeposited coating, it could not be stressed by pre-existing etched tracks. However, we note that even after 69 h of leaching, the tracks are strongly obstructed, which indicates that transfer of matter over small distances ($\sim 5,000$ Å) also occurs. For longer leach times (240 h), the tracks are slowly covered by small ($\sim 1,000$ Å) particles of unidentified nature and crystallinity which finally form a uniform surficial coating of granular aspect. These aspects support the conclusion that the inner part of the layer is indeed mainly formed *in situ*, whereas the external one, which grows slowly over long leach times, is mostly constituted of re-precipitated particles. Indeed, for more extreme experimental conditions, the formation of analcime or tobermorite, which can only be precipitated from the solution, has been confirmed^{8,11,18}.

In NaCl brine (250 g l^{-1} , Fig. 4), the hydrated layer is also composed of two sub-layers, the inner one showing a high density of micro-cracks ($0.1 \mu\text{m} \times 1 \mu\text{m}$) and a few very long fractures, and the external one being much smoother and highly fractured. In both sub-layers we observe clear remnants of the initial tracks, thus confirming that the hydrated layer is formed *in situ*. Moreover, the density of fractures in the external sub-layer, which are known to be due to the desiccation of silica gels, is directly related to the density of the tracks (Fig. 4a); indeed, most fractures are generated by the disruption of tracks (Fig. 4b).

We conclude that the hydrated layer is mostly formed *in situ* by a two-step mechanism and is not a purely re-precipitated coating. 'In situ' here means that the basic architecture of the layer must be derived from that of the pristine glass (for instance by hydration) if it is to 'keep the memory' of the presence of pre-etched tracks. However, its chemical and structural characteristics could also involve other complex phenomena, such as ion exchange, precipitation of diffusing species and polymerization of silica, as long as they imply the transport of small quantities of matter and/or very small distances compared with the characteristic dimensions of the tracks. However, in the case of one complex glass dissolved in H₂O at 100 °C, we have observed a probable re-precipitation step which produces the granules of the external sub-layer, thus confirming the observations of Nogues⁶.

We thank G. Baudin, M. Jorda and H. Bernas for their constant interest in this work, the French Atomic Energy Commission for financial support, N. Jacquet-Francillon for providing samples of complex glasses 3, 4 and 5, and M. Okuzumi (Jeol-France) for SEM micrographs.

Received 29 July; accepted 28 October 1985.

- Scholtze, H. *J. non-cryst. Solids* **52**, 91-103 (1982).
- Doremus, R. H. *J. non-cryst. Solids* **19**, 137-144 (1975).
- Lanford, W. A., Davis, K., Lamarche, P., Laursen, T. & Groleau, R. *J. non-cryst. Solids* **33**, 249-266 (1979).
- Smit, W., Stein, H. N. *J. non-cryst. Solids* **34**, 357-370 (1979).
- Smetts, B. M. J., Lommen, T. P. A. *Phys. Chem. Glasses* **24**, 35-36 (1983).
- Nogues, J.-L. thesis, Univ. Montpellier (1984).
- Crovisier, J.-L., Fritz, B., Grambow, B. & Eberhart, J.-P. *MRS Proc. Symp. Scientific Basis of Nuclear Waste Management, Stockholm, September 1985* (North-Holland, Amsterdam, in the press).
- Lutze, W., Malow, G., Ewing, R. C., Jercinovic, M. J. & Keil, K. *Nature* **314**, 252-255 (1985).
- Crovisier, J.-L., Thomassin, J.-H., Juteau, T. & Eberhart, J.-P. *Geochim. cosmochim. Acta* **47**, 377-387 (1983).
- Grambow, B. *MRS Proc.* **11**, 93-102 (1983).
- Bates, J. K., Jardine, L. J. & Steindler, M. J. *Science* **218**, 51-54 (1982).
- Thomassin, J.-H. thesis, Univ. Orléans (1984).
- Fleischer, R. L., Price, P. B. & Walker, R. M. *Nuclear Tracks in Solids* (University of California, 1975).
- Smetts, B. M. J., Lommen, T. P. A. *Phys. Chem. Glasses* **23**(3), 83-87 (1982).
- Della Mea, G., Dran, J.-C., Petit, J.-C., Bezzon, G. & Rossi-Alvarez, C. *Nucl. Instrum. Meth.* **218**, 493-499 (1983).
- Crank, F. *The Mathematics of Diffusion* (University of Oxford Press, 1978).
- Petit, J.-C., Langevin, Y., Lameille, J.-M. & Dran, J.-C. *MRS Proc.* **5**, 203-210 (1982).
- Lutze, W., Malow, G. & Rabe, H. *MRS Proc.* **15**, 37-45 (1983).
- Petit, J.-C. thesis, Univ. Paris, Orsay (1982).
- Arnold, G. W., Petit, J.-C. *Nucl. Instrum. Meth.* **209/210**, 1071-1077 (1983).

Pb–Pb systematics of Lewisian gneisses—implications for crustal differentiation

M. J. Whitehouse & S. Moorbath

Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK

Combined use of Rb–Sr, U–Pb and Sm–Nd decay schemes in geochronology provides insights into the evolution of ancient continental crust¹ such as the Lewisian of north-west Scotland. We report here a Pb–Pb whole-rock isochron date of $2,950 \pm 70$ Myr on amphibolite-facies grey gneisses from the northern region of the Lewisian; this is essentially the same as a Sm–Nd whole-rock isochron date ($2,920 \pm 50$ Myr) which was interpreted as the time of separation of igneous protoliths from the mantle². A period of ~ 240 Myr separates this date from the Scourian peak of granulite-facies metamorphism. This result indicates, for the Lewisian, that high-grade metamorphism and crustal stabilization cannot have been the result solely of thermal equilibration. More generally, it suggests that tectonic thickening was an important process during Archaean crustal evolution³ and that although whole-rock systems may remain open with respect to U–Pb until the termination of granulite-facies metamorphism, they may be closed during amphibolite-facies metamorphism.

The Lewisian Complex of mainland north-west Scotland is divided into three main regions which are considered to have originally occupied different crustal levels⁴. The central region, between Loch Laxford and Gruinard Bay, preserves granulite facies gneisses from the Scourian metamorphism which has been dated at $2,680 \pm 60$ Myr⁵. A recent estimate of metamorphic conditions indicates temperatures and pressures in excess of 1,000 °C and 8.5 kbar respectively⁶. Amphibolite facies gneisses are dominant in the northern and southern regions and previous isotopic work has shown that these record both the Scourian event with whole-rock Pb–Pb^{7–10} and zircon U–Pb¹¹ dates in the range 2,900–2,600 Myr as well as the Proterozoic Laxfordian event with Rb–Sr¹² and K–Ar¹³ dates in the range 1,880–1,550 Myr. Extensive granitic intrusions along the Laxford Front, the boundary between the northern and central regions, have been dated by Pb–Pb and Rb–Sr methods at $\sim 1,700$ Myr¹⁴.

A period of 100–200 Myr between extraction of igneous precursors of crustal gneisses and their subsequent stabilization has been recognized in many Archaean regions from studies of initial ⁸⁷Sr/⁸⁶Sr ratios^{1,15}. Stabilization is produced by high-grade metamorphism purging the lower crust of heat-producing elements (K, Rb, U and Th) and other large-ion lithophiles (LIL)¹⁶. The whole process from accretion to final differentiation has been termed a 'crustal accretion-differentiation superevent' (CADS)^{1,15}. The Sm–Nd date of $2,920 \pm 50$ Myr reported by

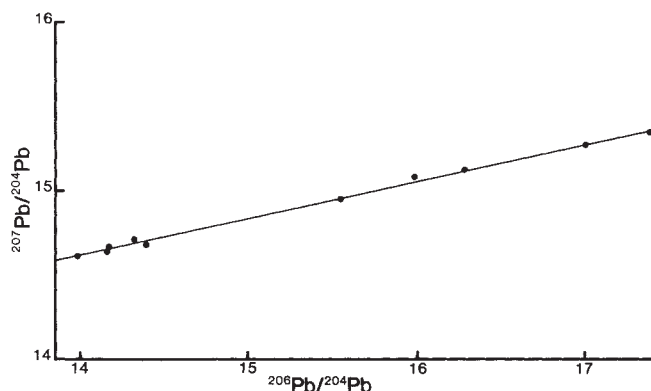


Fig. 1 $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ isochron plot of the grey gneiss data. Age = $2,950 \pm 70$ Myr; $\mu_1 = 7.82$; MSWD = 1.94.

Hamilton *et al.*² was obtained on a lithologically diverse suite of gneiss samples, which are not necessarily co-genetic, from the northern and central regions. The difference of ~ 240 Myr between this and Scourian dates obtained by other methods was interpreted² as the approximate duration of the Scourian CADS.

North of the Laxford Front the distribution of Laxfordian intrusions and migmatization becomes more localized, most of the region consisting of gneisses belonging to the three intergradational petrographic divisions recognized by the Geological Survey¹⁷—that is, 'grey gneiss', in which biotite is the principal ferromagnesian mineral, 'hornblende-biotite gneiss' and 'dark gneiss', in which ferromagnesian minerals predominate. Samples collected for the present study were chosen so as to minimize the effects of post-Scourian disturbance. Gneisses containing, or adjacent to, discordant granitic and pegmatitic intrusions, as well as finely banded gneisses, which may have been affected by fluid movement, were avoided. Sample localities are given in Table 1.

Pb was extracted using standard dissolution and ion-exchange techniques. Details of Pb mass spectrometry are given in ref. 14. Decay constants for ²³⁸U and ²³⁵U were taken as $1.55125 \times 10^{-10} \text{ yr}^{-1}$ and $9.85484 \times 10^{-10} \text{ yr}^{-1}$ respectively. All dates quoted here from early publications have been recalculated using these values, and are quoted with a 2σ error. The analyses are presented in Table 1 and $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ ratios are plotted in Fig. 1.

A 10-point isochron (MSWD = 1.94) yields a date of $2,950 \pm 70$ Myr (2σ). A model $^{238}\text{U}/^{204}\text{Pb}$ (μ_1) ratio of 7.82 for two-stage isotopic evolution is typical of late Archaean upper mantle source regions¹.

This result indicates that the whole-rock U–Pb system in the grey gneisses has remained undisturbed since late Archaean times despite significant Laxfordian reworking at $\sim 1,700$ Myr.

Table 1 Analytical data

Sample	Locality	Grid ref.	$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$
MJW192	Rhiconich	NC251534	14.399	14.676	36.578
MJW193	Rhiconich	NC251534	13.986	14.612	34.992
MJW220	Rhiconich	NC252538	14.328	14.711	36.372
MJW211	Kinlochbervie	NC230563	14.177	14.668	36.071
MJW241	Laxford	NC233485	14.171	14.641	34.799
MJW251	Rispond	NC449656	17.380	15.337	38.699
MJW252	Rispond	NC449656	16.285	15.117	36.533
MJW253	Rispond	NC449656	16.997	15.260	38.336
MJW254	Rispond	NC449656	15.983	15.075	35.396
MJW255	Rispond	NC445653	15.548	14.951	36.624

Within-run precision of Pb isotope ratios between 0.02 and 0.05% (2σ). Analyses are corrected for mass fractionation of $-0.1 \pm 0.03\%$ per AMU. Overall analytical errors $< 0.1\%$.

The measured date is similar to the Sm-Nd accretion date of $2,920 \pm 50$ Myr² but outside the error of published Scourian Pb-Pb and Rb-Sr dates. This result indicates that closure of this whole-rock system in the northern region, shortly after crystallization, occurred over 200 Myr earlier than the peak of high-grade metamorphism in the central region. Any model for early evolution of the Lewisian crust must provide for granulite-facies conditions in the lower crust after this lengthy time interval. In addition, the preservation of an older date from amphibolite-facies gneisses suggests that although final closure of whole-rock U-Pb systems occurs only on cessation of granulite-facies metamorphism, these systems may remain closed during amphibolite-facies and lower grades of metamorphism.

The thermal history of magmatically accreted crust has been considered by Wells¹⁸ for three accretion scenarios—under-accretion in which magma is plated onto the base of existing crust; over-accretion in which magma is emplaced at a high crustal level; and random accretion in which the level of emplacement within the crust varies randomly. The over-accretion and random-accretion models can produce temperature maxima in the lower crust after cessation of magma addition, though this occurs only for accretion events of less than ~35 Myr duration and cannot explain a time interval as long as 240 Myr. Under-accretion cannot produce a post-magmatic temperature maximum.

An alternative to thermal models which rely only on the conductive heating effect of newly emplaced intrusions as well as radioactive heating, is one involving tectonic thickening of unstable juvenile crust³. The thermal model considered by Hamilton *et al.*² to explain the length (~240 Myr) of the observed time interval assumed constant crustal thickness and implied granulite *P-T* conditions and extensive melting by ~2,820 Myr, unless the heat-producing and LIL elements migrated to higher structural levels before ~2,820 Myr. This is inconsistent with Sr-isotope data (M.J.W., unpublished) which indicate that the average ⁸⁷Rb/⁸⁶Sr ratio between crustal accretion at ~2,920 Myr and high-grade metamorphism at ~2,680 Myr was high and shows no evidence for Rb depletion. Oxburgh and Turcotte¹⁹ have discussed the perturbation of the regional conductive geothermal gradient caused by overthrusting with specific reference to the Eastern Alps. Once an equilibrium gradient has been restored the overall effect is one of increased metamorphic grade in the lowermost tectonic units. The style of the predominantly flat-lying *f*₁ foliation in the Lewisian²⁰ is consistent with major horizontal movements associated with thrusting. This model could also account for occurrence of gneisses of supracrustal origin throughout the complex²¹.

The juxtaposition of crustal levels across the Laxford Front as well as the geochronological data reported here require a major thermal event long after an equilibrium conductive geothermal gradient has been established. Large-scale intra-continental thrusting and associated crustal thickening is one mechanism for producing the requisite high-grade metamorphism which terminated the Scourian CADS at ~2,700 Myr.

We thank P. N. Taylor and R. St J. Lambert for constructive criticisms. M.J.W. acknowledges receipt of a NERC studentship.

Received 1 August; accepted 12 November 1985.

1. Moorbath, S. & Taylor, P. N. in *Precambrian Plate Tectonics* 491–525 (ed. Kröner, A.) (Elsevier, Amsterdam, 1981).
2. Hamilton, P. J., Evensen, N. M., O'Nions, R. K. & Tarney, J. *Nature* **277**, 25–28 (1979).
3. Kröner, A. *Rev. Bras. Geocienc.* **12**, 15 (1982).
4. Sheraton, J. W., Skinner, A. C. & Tarney, J. in *The Precambrian of Scotland and Related Rocks of Greenland* 13 (eds Park, R. G. & Tarney, J.) (Department of Geology, University of Keele, 1973).
5. Chapman, H. J. & Moorbath, S. *Nature* **268**, 41–42 (1978).
6. Barmicoat, A. C. *J. metamorphic Geol.* **1**, 163–182 (1983).
7. Moorbath, S. & Park, R. G. *Scott. J. Geol.* **8**, 51–74 (1972).
8. Moorbath, S., Powell, J. L. & Taylor, P. N. *J. geol. Soc.* **131**, 213–222 (1975).
9. Moorbath, S., Welke, H. & Gale, N. H. *Earth planet. Sci. Lett.* **6**, 245–256 (1969).
10. Holland, J. G. & Lambert, R. St J. *Precamb. Res.* **2**, 161–188 (1975).
11. Lyon, T. D. B. & Bowes, D. R. *Krystallinikum* **13**, 53–72 (1977).
12. Chapman, H. J. thesis, Univ. Oxford (1978).

13. Lambert, R. St J. & Holland, J. G. *J. geol. Soc.* **128**, 3–19 (1972).
14. Taylor, P. N., Jones, N. W. & Moorbath, S. *Phil. Trans. R. Soc. A310*, 605–625 (1984).
15. Moorbath, S. *Chem. Geol.* **20**, 151–187 (1977).
16. Tarney, J. & Windley, B. F. *J. geol. Soc.* **134**, 153–172 (1977).
17. Peach, B. N. *et al. Mem. geol. Surv. UK* **33**, 268 (1907).
18. Wells, P. R. A. *Earth planet. Sci. Lett.* **46**, 253–265 (1980).
19. Oxburgh, E. R. & Turcotte, D. L. *Schweiz. miner. petrogr. Mitt.* **54**, 641–662 (1974).
20. Sheraton, J. W., Tarney, J., Wheatley, T. J. & Wright, A. E. in *The Precambrian of Scotland and Related Rocks of Greenland*, 31 (eds Park, R. G. & Tarney, J.) (Department of Geology, University of Keele, 1973).
21. Watson, J. *Spec. Rep. geol. Soc. Lond.* **6**, 15–29 (1975).

Stromatolites from the 3,300–3,500-Myr Swaziland Supergroup, Barberton Mountain Land, South Africa

Gary R. Byerly, Donald R. Lower & Maud M. Walsh

Department of Geology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

A morphologically variable assemblage of stromatolites has been discovered in thin chert layers within the Fig Tree Group of the Swaziland Supergroup, South Africa. They are commonly low-relief, nearly stratiform, laterally linked domes. Rarer forms include pseudocolumns and crinkly stratiform stromatolites. The stromatolites grew on a substrate of altered komatiitic lava and sediments deposited on the lava surface, and in most places are covered by later komatiitic flows. Abundant fine-grained tourmaline included within the stromatolite laminae suggests that stromatolites formed in an environment dominated by boron-rich hot-spring emissions and evaporitic brines.

The Swaziland Supergroup of the Barberton Mountain Land consists of a lower sequence (Onverwacht Group) of ultramafic and mafic lava flows containing thin interbedded cherty layers, an overlying sequence of chert, clastic sediments, and some interbedded volcanics (Fig Tree Group), and an uppermost unit of primarily conglomerate, sandstone and shale (Moodies Group). A minimum age of 3,300 Myr for the Fig Tree Group is provided by K-Ar dating of low-grade regional metamorphism¹, and the lower part of the Onverwacht is considered to have an age of 3,500 Myr based on Nd-Sm techniques². Stromatolites occur in three separate areas (Fig. 1). The most extensive development is at locality A where they are found throughout an area roughly 1 × 2 km in an isoclinal antiform broken up by small throw, high-angle faults.

At localities A and B, the stromatolites overlie silicified komatiite flows showing well-developed spinifex and cumulus textures. The surfaces of the flows are irregular, however, and in places there is from 1 to 18 m of sedimentary rocks between the altered lava and the stromatolitic cherts. These associated sediments include banded ferruginous and carbonaceous chert; thin, graded layers of volcanic accretionary lapilli, ash and dust; and minor sandstones derived by erosion of the underlying volcanic rock. Also common are fine breccias containing debris eroded from the stromatolites. The stromatolites were covered by later komatiitic lava flows up to 20 m thick. Most are massive but some include pillowed or hyaloclastite facies. Spinifex-textures are widely developed. The bases of these flows commonly form casts of the stromatolite surface. At least one other stromatolitic chert unit is interbedded with the later komatiitic lavas at locality A.

At locality B the stromatolites grew directly on the top of an irregular pinnacle of a silicified komatiitic flow rock. Laterally, banded chert, detrital sediments including stromatolitic debris, and volcanoclastic layers overlie the silicified volcanic basement. The stromatolitic zone is succeeded by >300 m of cherts and clastic sediments of the Fig Tree Group. At location C the stromatolites are in a tectonically isolated block that is surrounded by komatiites. The stromatolites are underlain by 1–2 m of

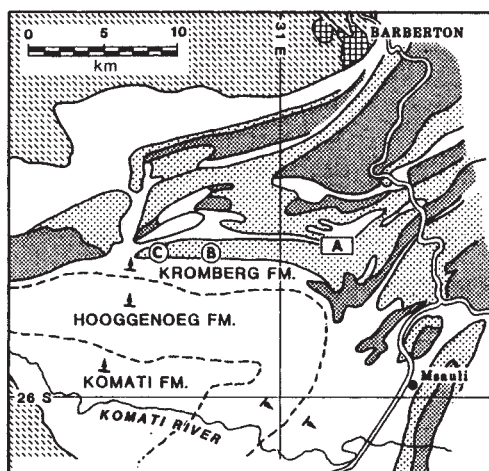


Fig. 1 Map of the Barberton Greenstone Belt with newly discovered stromatolite localities A, B and C in the Fig Tree Group. No pattern, Onverwacht Group; coarse stipple, Fig Tree Group; fine stipple, Moodies Group; and diagonal pattern, granitic rocks. Locality A is a prominent chert ridge on the south-east side of the narrow valley at $25^{\circ}54'10.5''$ S and $31^{\circ}02'41''$ E; locality B is a smaller outcrop on the south-facing ridge about 0.5 km north of the Msauli River at $25^{\circ}54'28''$ S and $30^{\circ}55'33''$ E; locality C is a very small outcrop at $25^{\circ}54'15.5''$ S and $30^{\circ}53'40.3''$ E along an old forestry road.

banded black and white chert and may be overlain by lavas. At locality A, the only site where the stromatolites can be traced laterally, the stromatolitic layers are thin and laterally discontinuous, ranging from <1 to ~ 20 cm thick and typically extending for about 5 m along strike. They pinch out over highs on the underlying lavas. Individual chert units interbedded with komatiitic lava flows commonly include two or three stromatolite layers separated by 10–40 cm of non-stromatolitic sediment.

The most common stromatolites are low-relief domes ranging from ~ 1 cm wide and 0.5 cm high to ~ 3 cm wide and 3 cm high. In some places, the small domes have developed upwards into larger domes, compound domes, or pseudocolumns with heights up to 10 cm (Figs 2 and 3). The large domes generally have gently convex, almost flat crests, but some are concave. Both the large and small domes are slightly asymmetrical, and plan views show some elongation of domes. The domes are, in places, associated with stratiform stromatolites with fine crinkly laminae. Synoptic relief of laminae ranges from a few millimetres in the small domes to ~ 1.5 cm in the larger domes, with greater degrees of inheritance in the larger domes and pseudocolumns. In some places, laminae form bridges between adjacent domes (Fig. 2).

In all areas, the stromatolites are preserved within grey to black chert that shows fine laminae consisting mainly of microcrystalline quartz. Both light and dark laminae range in thickness from 50 to 100 μm and are made up primarily of microcrystalline quartz of 2–4 μm grain size. The darker laminae contain very fine, submicrometre-sized material that is possibly organic matter. In many samples, the dark laminae also include tourmaline crystals ranging from 25 to 200 μm , and 25 to 50 μm crystals of sericite, haematite and rutile. Individual laminae may be traced through single tourmaline crystals. Composite laminae commonly up to 500 μm in thickness can be recognized by their variable tourmaline or sericite contents. Some composite laminae contain up to 50% tourmaline. Fine sandy layers of stromatolite fragments occur between some laminae.

Stromatolites are 'organosedimentary structures produced by sediment trapping, binding and/or precipitation as a result of the growth and metabolic activity of microorganisms'. The identification of the Fig Tree structures as stromatolites is based



Fig. 2 Stromatolites directly overlying brecciated top of komatiitic lava flow. Small basal domes are covered by large, slightly asymmetrical domes or pseudocolumns. These are, in turn, overlain by very low-relief laminae with high inheritance. Note also the stromatolite flake breccia in the lower layer extending down into black chert surrounding komatiitic clasts. Samples have been slabbbed and polished, then etched in hydrofluoric acid to bring out the fine detail of laminae. Coin diameter, 2.5 cm.



Fig. 3 Pseudocolumnar stromatolites in a thin (1-m) chert between komatiitic lava flows at locality C. Columns are inclined, and composed of wrinkled laminae with high inheritance.

on their similarity to previously described stromatolites. Both the domal and pseudocolumnar macrostructures and the fine crinkly laminations argue for a biological influence on the sediment. The Fig Tree stromatolites most notably resemble both stratiform and columnar stromatolites of the Proterozoic Gunflint Iron Formation^{4–6}. Stromatolite-like structures may be formed inorganically, but the structures reported here lack obvious features, such as cross-lamination or discordant relationship to other sediments, that would indicate an evaporitic or other non-biological origin⁷.

It has been suggested that early Archaean stromatolite-building organisms were not unlike those of the Proterozoic and Phanerozoic: probably filamentous, but possibly coccoidal bacteria or cyanobacteria^{8,9}. Microfossils have not been found within the stromatolitic rocks of the Fig Tree, but coccoidal microfossils have been described in black cherts of the lower Fig Tree¹⁰ and filamentous microfossils found in the immediately underlying Kromberg Formation^{11,12}.

The morphological characteristics of stromatolites probably reflect both the type of organism that constructed the stromatolite and environmental factors⁶. Certain features of the Fig Tree stromatolites have environmental implications. The predominance of lateral linkage between both stromatolite domes and low-relief mounds suggests that wave or current action was slight enough that intermound growth was undisturbed¹³. The stromatolites formed mainly on basement highs, indicating that one or more of the following might have been required for stromatolite development: shallow water, availability of a hard substrate, or sparse to absent clastic sedimentation. The common occurrence of stromatolite flake breccias suggests that subaerial exposure and desiccation of stromatolites occurred, although storms might have eroded subaqueous stromatolites. In areas away from the actual growth sites, the presence of stromatolite flake debris may provide a useful indicator of the existence of stromatolites.

The tourmaline that is commonly found within the stromatolitic laminae suggests the influence of silica, magnesia, iron- and boron-enriched waters¹⁴. Tourmaline-rich sedimentary layers are associated with Proterozoic stratabound sulphide deposits considered to be of marine exhalative origin and from evaporitic environments where they are associated with carbonate sediments¹⁴.

The stromatolites reported here are widespread, morphologically variable, and associated with lithologically diverse sedimentary units that include a possible evaporitic component. The other two occurrences of well-documented early Archaean stromatolites are in the Warrawoona Group of Western Australia. In this area, the regionally extensive Strelley Pool Chert⁸

includes laterally linked conical pseudocolumnar stromatolites interbedded with silicified evaporites, and the Towers Formation locally contains stromatolites⁹ similar to those described here, also interbedded with deposits interpreted to represent evaporites. The results of these studies suggest that early Archaean stromatolites may have been widely developed within a range of shallow-water settings, but may have been particularly favoured by conditions associated with near-evaporitic or evaporitic environments.

Field studies in the Barberton Mountain Land were supported by NSF grants EAR7919907 to G.R.B. and D.R.L., EAR7919908 and EAR8206015 to D.R.L., and an NSF Fellowship to M.M.W. The ARCO Foundation, Anglo-American Corporation, and the University of Cape Town Departments of Geochemistry and Geology provided additional support.

Received 19 April; accepted 11 October 1985.

1. deWit, M. J., Hart, R., Martin, A. & Abbot, P. *Econ. Geol.* **77**, 1783-1802 (1982).
2. Hamilton, P. J., Evensen, M. M., O'Nions, R. K., Smith, H. S. & Erlank, A. *J. Nature* **297**, 298-300 (1979).
3. Awramik, S. M. & Margulis, L. in *Stromatolites* (ed. Walter, M. R.) 1 (Elsevier, Amsterdam, 1976).
4. Awramik, S. M. in *Stromatolites* (ed. Walter, M. R.) 311-336 (Elsevier, Amsterdam, 1976).
5. Awramik, S. M. & Semikhatov, M. A. *Can. J. Earth Sci.* **16**, 484-495 (1979).
6. Semikhatov, M. A., Gebelein, C. D., Cloud, P., Awramik, S. M. & Benmoire, W. C. *Can. J. Earth Sci.* **16**, 992-1015 (1979).
7. Walter, M. R. in *Earth's Earliest Biosphere* (ed. Schopf, J. W.) 187-213 (Princeton University Press, 1983).
8. Lowe, D. R. *Nature* **284**, 441-443 (1980).
9. Walter, M. R., Buick, R. & Dunlop, J. S. R. *Nature* **284**, 443-445 (1980).
10. Knoll, A. H. & Baughnorn, E. *S. Science* **198**, 396-398 (1977).
11. Walsh, M. M. & Lowe, D. R. *Abstr. 14th lunar Planet. Sci. Conf.* **2**, 814 (1983).
12. Walsh, M. M. & Lowe, D. R. *Nature* **314**, 530-532 (1985).
13. Logan, B. W., Rezak, R. & Ginsburg, R. N. *J. Geol.* **72**, 68-83 (1964).
14. Slack, J. F., Herriman, N., Barnes, R. C. & Plimer, I. R. *Geology* **12**, 713-716 (1984).

Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms

**J. S. Wainscoat, A. V. S. Hill, A. L. Boyce*, J. Flint,
M. Hernandez, S. L. Thein, J. M. Old, J. R. Lynch,
A. G. Falusi†, D. J. Weatherall & J. B. Clegg**

Medical Research Council Molecular Haematology Unit,
Nuffield Department of Clinical Medicine, University of Oxford,
John Radcliffe Hospital, Oxford OX3 9DU, UK

* Department of Biological Anthropology, University of Oxford,
Oxford OX2 6QS, UK

† Department of Haematology, University College Hospital, Ibadan, Nigeria

The genetic relationships of human populations have been studied by comparing gene frequency data for protein and blood-group loci of different populations^{1,2}. DNA analysis now promises to be more informative since not only do the DNA coding sequences have more variation than their corresponding proteins but, in addition, noncoding DNA sequences display more extensive polymorphism^{3,4}. We have now studied the frequency of a group of closely linked nuclear DNA polymorphisms (haplotypes)⁵ in the β -globin gene cluster of normal (β^A) chromosomes of individuals from eight diverse populations. We have found that all non-African populations share a limited number of common haplotypes whereas Africans have predominantly a different haplotype not found in other populations. Genetic distance analysis based on these nuclear DNA polymorphisms indicates a major division of human populations into an African and a Eurasian group.

Genetic distance analysis for protein loci suggests that during human evolution the negroid group and the caucasoid-mongoloid group diverged first and that the caucasoid and mongoloid groups diverged from each other at a later date². In

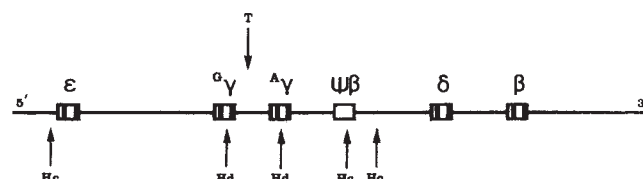


Fig. 1 Schematic representation of the β -globin gene cluster; the linkage of the five polymorphic restriction enzyme sites shown below the cluster constitutes the haplotype studied. Hc, *HincII*; Hd, *HindIII*. The polymorphic *TaqI* site (T) is shown above the cluster. DNA was recovered from peripheral blood by phenol-chloroform extraction, and restriction enzyme analysis performed as described in ref. 24. The haplotypes were determined from the *HincII* and *HindIII* digests as in ref. 5. The *TaqI* polymorphic restriction site is present in all of 100 representative non-African chromosomes studied, resulting in $\gamma\gamma$ restriction-fragment lengths of 1.7 and 2.5 kilobases. It is absent in 10 African ----- haplotypes, giving $\gamma\gamma$ restriction fragment lengths of 1.7 and 3.0 kilobases.

contrast, phylogenetic trees of human races based on blood groups⁶ and histocompatibility locus antigens⁷ have their primary divisions between the mongoloid and the caucasoid-negroid groups. It has been considered that the protein data may be more reliable than blood-group results because they more closely reflect variation in DNA sequence. However, the best way to study population affinities is by comparing DNA sequences of individuals from different populations, since not only is DNA analysis more informative than analysis of proteins but it is also direct and unambiguous. Human restriction enzyme polymorphisms now provide a method of detecting DNA sequence changes and can be used to examine large numbers of individuals. A previous study of mitochondrial DNA polymorphisms in different populations indicated that the two African populations formed a distinct lineage from the others studied⁸. This observation was interpreted as being due to a higher mutation rate in the Bushman population, although there

Table 1 Numbers of chromosomes possessing the various haplotypes in each population

Haplotypes	British	Cypriot	Italian	Asian Indian	Thai	Melanesian A	Melanesian B	Polynesian	African A	African B
+-----	16	59	32	58	29	44	73	43	0	3
-+---+	15	14	15	28	0	10	19	6	2	4
-++-+	5	8	2	15	2	3	0	0	1	0
+---++	0	0	0	0	0	3	8	0	0	0
-++---	1	1	1	3	0	0	0	0	0	0
-++++	0	0	0	0	0	0	5	1	0	0
++-++	0	0	0	3	0	0	0	0	0	0
-----	0	0	0	1	0	0	2	1	0	0
++-++	0	0	0	1	0	0	0	0	0	0
---++	0	0	0	1	1	0	0	0	0	0
++----	0	0	0	1	0	0	0	0	0	0
-+----	0	0	0	0	0	0	0	0	1	1
-+---+	0	0	0	0	0	4	0	4	6	6
-----+	0	0	0	0	0	2	0	0	16	21
Totals	37	82	50	111	32	66	107	55	26	35

The Italian, Cypriot and Asian Indians studied were parents of thalassaemic children referred for genetic analysis. The haplotypes of the normal (β^A) chromosomes were determined by family linkage studies. All the other population groups consisted of healthy individuals with normal haemoglobin electrophoresis results. The first group of Africans (African A) were university students originating from sub-Saharan countries: Gambia, Kenya and Nigeria; the second group of Africans (African B) were Nigerians mainly from the Yoruba, Bendel and Ibo. The Melanesians and Polynesians studied were individuals showing normal haemoglobin electrophoresis results taken from a larger study of the incidence of the haemoglobinopathies in those populations. The first group of Melanesians was from the eastern and southern highlands of Papua New Guinea; the second group was from north coastal Papua New Guinea and Vanuatu. The Thai individuals were from Bangkok and the British individuals were from Oxford. All first-degree relatives have been excluded from the study. A polymorphic restriction site is denoted as being present in a particular haplotype by a + symbol and absent by a - symbol.

is no direct evidence for this. Mitochondrial DNA has a maternal pattern of inheritance, which means that the dispersal rate of mitochondrial genes is that of females rather than being representative of both sexes. We have therefore sought to investigate human population relationships by analysing nuclear DNA polymorphisms and comparing these data with the mtDNA polymorphism results.

We have studied five linked polymorphic restriction enzyme sites in the β -globin gene cluster⁵ (Fig. 1), the combination of the sites along one chromosome being referred to as a haplotype. The five sites show strong linkage disequilibrium and are 5' to a recombination hotspot in the gene cluster⁵. Although these restriction sites are in a functional gene cluster, we have analysed only those chromosomes bearing normal β^A -globin genes on which there is no known selection pressure. The numbers of the haplotypes found are reported for eight population groups; British, Italian, Cypriot, Asian Indian, Melanesian, Polynesian, Thai and African. Analysis of 601 chromosomes revealed 14 of the 32 (2⁵) possible haplotypes (Table 1). One haplotype (+-----) is the most common in all seven non-African populations and is found at similar frequencies in each of them. The second most common haplotype (-+---+) is found in all the non-African populations with the exception of the Thai population, although it has been observed previously in South-East Asia⁹. A third haplotype (-++-+) accounts for the majority

of the remaining haplotypes found in the non-African populations. The two African populations have a common haplotype (-----+) which is absent in all the non-African populations, with the exception of two individuals from the Papua New Guinea (PNG) highland population. However, even these two exceptions may be distinguished from the common African haplotype by the presence of a *TaqI* restriction-site polymorphism 3' to the γ gene (Fig. 1). Hence it seems probable that this PNG haplotype has an independent non-African origin through recombination. Two other new haplotypes (+---++ and -++++), presumably formed by recombination events, were found in the Melanesian and Polynesian populations.

It is interesting that none of these four common haplotypes can be derived from any of the other three by either a single crossover or by a single-base mutation causing the loss or acquisition of a restriction enzyme site. We suggest, therefore, that these four haplotypes predate the racial divergence and hence that these structures have been stable for long periods of time of the order of 100,000 yr. For example, the highland Melanesian population has similar haplotype frequencies to the Europeans from which they have been separated for 20,000–40,000 yr¹⁰. Most of the rarer haplotypes may be derived from the four common ones by single crossovers. Theoretically there are 13 haplotypes which may be accounted for in this way, and 8 have been observed in our study, in contrast to only 3 of a possible

Table 2 Genetic distances for the 10 populations

	Balakrishnan-Sanghvi (G_c) distances							
	British	Cypriot	Italian	Thai	Indian	Melanesian	Polynesian	African
British	-							
Cypriot	0.6908	-						
Italian	0.5363	0.4012	-					
Thai	1.3064	0.7366	0.9695	-				
Indian	0.6472	0.6533	0.6869	1.0734	-			
Melanesian	1.0239	0.6843	0.6977	0.9533	0.9872	-		
Polynesian	1.1368	0.6650	0.7323	0.8305	1.0360	0.5975	-	
African	2.8658	2.8399	2.8300	2.9619	2.8719	2.7742	2.7288	-

The haplotype data of the two Melanesian populations have been combined, as have the haplotype data of the two African populations.

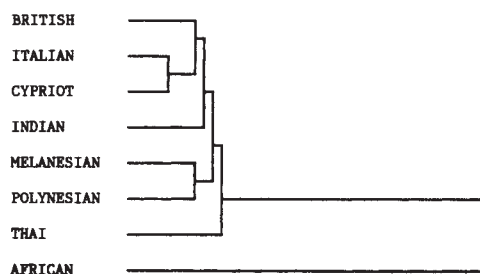


Fig. 2 Dendrogram for the eight population groups studied.

remaining 15 haplotypes which cannot be accounted for by single crossovers between the common haplotypes.

Genetic distance analysis of the populations was undertaken using the G_c distance measure of Balakrishnan and Sanghvi¹¹ (Table 2). The dendrogram based on these distances, using single-linkage cluster analysis¹², is shown in Fig. 2. Certain population groups which are close geographically are separated by small genetic distances, for example, the Cypriot and Italian, and the Polynesian and Melanesian populations. However, a striking finding is the large genetic distance observed between the African and non-African populations. Interestingly, among the known race-specific polymorphisms there does seem to be a predominance of unique alleles at both protein and blood-group loci in negroid populations¹³. Furthermore, several negroid-specific DNA polymorphisms have been described at well-characterized loci, for example, α -globin¹⁴, β -globin^{15,16}, growth hormone¹⁷, dihydrofolate reductase¹⁸ and insulin¹⁹. Similarly, a negroid-specific haplotype has been described at the albumin locus²⁰. These nuclear DNA data seem to support a primary division of human populations into an African and a Eurasian group. Thus, it may not be necessary to invoke a difference in mutation rate between populations to explain the genetic distance pattern observed by mtDNA analysis.

The earliest fossils of anatomically modern man (*Homo sapiens sapiens*) have been found in Africa at Omo in Ethiopia, Border Cave in South Africa and at Klasies River Mouth in South Africa²¹. The data from the last site suggest that *H. sapiens sapiens* was present in South Africa more than 100,000 yr ago²², and an adult mandible from Border Cave has been dated to about 90,000 yr BP²³. Hence, it has been argued that the evolution of modern man took place in Africa²¹. Our data are consistent with such a scheme, in which a founder population migrated from Africa and subsequently gave rise to all non-African populations. Such a group would appear to have lost the now predominant African haplotype (----+), suggesting that the founder population was small. Comparative studies of other groups of nuclear DNA polymorphisms, particularly those which are population-specific, should provide further information on the prehistoric migrations and genetic affinities of human populations.

We thank Professors P. Wasi and G. J. F. Essan and Drs M. Sampietro, D. K. Bowden, S. J. Oppenheimer, R. J. Trent, K. Mickleson, K. Bhatia and B. Modell for help with the collection of samples for the original studies, and the Rockefeller Foundation for financial support.

Received 2 September; accepted 9 December 1985.

1. Nei, M. & Roychoudhury, A. K. *Science* **177**, 434-436 (1972).
2. Nei, M. & Roychoudhury, A. K. *Evol. Biol.* **14**, 1-59 (1982).
3. Jeffreys, A. J. *Cell* **18**, 1-10 (1979).
4. Cooper, D. N., Smith, B. A., Cooke, H. J., Niermann, S. & Schmidtke, J. *Hum. Genet.* **69**, 201-205 (1985).
5. Antonarakis, S. E., Boehm, C. D., Giardina, P. J. V. & Kazazian, H. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 137-141 (1982).
6. Cavalli-Sforza, L. L. & Edwards, A. W. F. *Genetics Today* (Pergamon, Oxford, 1964).
7. Piazza, A., Sgarbetta-Zonta, L., Gluckman, P. & Cavalli-Sforza, L. L. *Tissue Antigens* **5**(5), 445-463 (1975).
8. Johnson, M. J., Wallace, D. C., Ferris, S. D., Rattazzi, M. C. & Cavalli-Sforza, L. L. *J. molec. Evol.* **19**, 255-271 (1983).

9. Cheng, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2821-2825 (1984).
10. Bellwood, P. *Man's Conquest of the Pacific* (Collins, London, 1978).
11. Balakrishnan, V. & Sanghvi, L. D. *Biometrics* **24**, 859-865 (1968).
12. Sneath, P. H. A. & Sokal, R. R. *Numerical Taxonomy* (Freeman, San Francisco, 1973).
13. Mourant, A. E. *Histocompatibility Testing*, 31-37 (1972).
14. Goodbourn, S. E. Y., Higgs, D. R., Clegg, J. B. & Weatherall, D. J. *Molec. Biol. Med.* **2**, 223-238 (1984).
15. Kan, Y. W. & Dozy, A. M. *Science* **209**, 492-493 (1980).
16. Orkin, S. H., Antonarakis, S. E. & Kazazian, H. H. *Prog. Haemat.* **13**, 49-73 (1983).
17. Chakravarti, A., Phillips, J. A., Mellits, K. H., Buetow, K. H. & Seeburg, P. H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6085-6089 (1984).
18. Anagnou, N. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5170-5174 (1984).
19. Bell, G. I., Horita, S. & Karam, J. H. *Diabetes* **33**, 176-183 (1984).
20. Murray, J. C., Mills, K. A., Demopoulos, C. M., Horung, S. & Motulsky, A. G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3486-3490 (1984).
21. Stringer, C. in *Hominid Evolution* (ed. Foley, R.) 55-83 (Academic, New York, 1984).
22. Singer, R. & Wymer, J. *The Middle Stone Age at Klasies River Mouth in South Africa* (University of Chicago Press, 1982).
23. De Villiers, H. S. *Afr. J. Sci.* **72**, 213-215 (1976).
24. Old, J. M. & Higgs, D. R. *Meith. Haemat.* **6**, 74-102 (1982).

Crabs that breathe air with their legs—*Scopimera* and *Dotilla*

David P. Maitland

School of Zoology, University of New South Wales, PO Box 1, Kensington, New South Wales, Australia 2033

Sand-bubbler crabs, *Scopimera* and *Dotilla* (Crustacea, Brachyura, Ocypodidae, Scopimerinae), are small (carapace width 1 cm) round-bodied crabs which occur in vast numbers throughout the Indo-Pacific on suitable tropical and sub-tropical sheltered sandy beaches. At high tide they are found in burrows that contain a trapped pocket of air in which the crabs are able to breathe until low tide. During low tide they become active and emerge from their burrows to feed on surface detritus and stranded plankton¹⁻³. One of the most striking, and unique, features of these crabs is the presence of large membranous disks ('tympana') on the meral segments of the legs (Fig. 1) and sometimes (in *Dotilla*) on the thoracic sterna³. Although in 1893 Aurivillius suggested that the membranes might be used for hearing⁴, their function has remained unknown. Here I present evidence which suggests that, rather than being auditory organs, the tympana are primarily respiratory structures used in aerial gas exchange. As the generic name *Scopimera* means "thighs with windows in them"⁵, I have coined the term 'gas windows' for these membranes.

Scopimera inflata, which is endemic to Australia, is a typical scopimerine crab with slender legs and meri expanded and flattened in the vertical plane. On both sides of each merus (including the chelae) is a large oval membrane stretched across the curved frame of the leg—the gas window (see Fig. 2A). Immediately below the gas-window membrane lies a complex system of blood spaces (Fig. 2B-D). To visualize these spaces, autotomized legs were injected with Batson's No. 17 corrosion compound (Polysciences, Warrington, Pennsylvania) into the leg sinus between the dactylus and propodus joint (see ref. 6 for additional methods).

In general, a single artery supplies the legs of crabs with oxygenated blood and venous blood drains into the leg sinus which surrounds the muscles like a sleeve⁷. In *S. inflata* the venous blood flows relatively freely back up the leg sinus towards the body until it reaches the meral segment; here it is divided and collected in a complex series of large blood channels formed by a matrix of connective tissue sheets and partitions (Fig. 2C). The blood is directed towards the respiratory epithelium within branching channels of ever decreasing diameter (30 μ m to 15 μ m to 5 μ m) until it is spread as a thin layer (3-5 μ m) directly below the gas-window membrane (Fig. 2D) (compare with the lacunar beds of crab lungs^{6,8}). The blood from the smaller spaces ('lacunae') is subsequently guided away from the membrane within a closely interdigitating series of larger, flattened channels

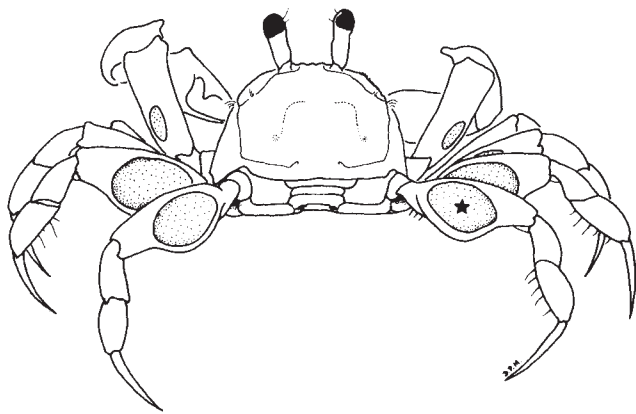


Fig. 1 The sand-bubbler crab *Scopimera inflata* is typical among scopimerine crabs in possessing a pair of membranous disks (gas windows) in each leg meral segment (indicated by a star). The crab measures 1 cm across the carapace.

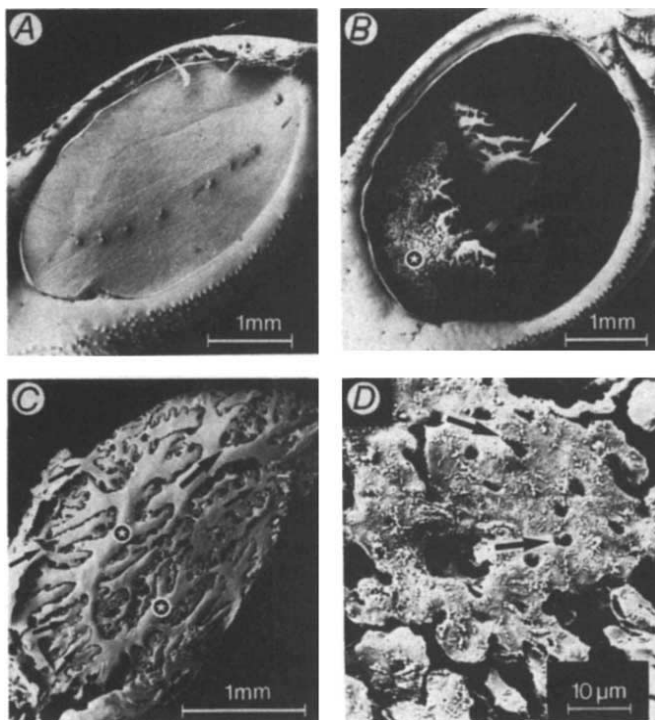


Fig. 2 **A**, Scanning electron micrograph (SEM) of the anterior gas window on the merus of the second right walking leg of *S. inflata* (dorsal is top left and proximal is top right). **B–D**, SEMs of plastic casts of gas-window blood channels in the legs of *S. inflata*. Blood flows towards the body from lower left to upper right in each case. **B**, A partial cast showing the organization of blood channels in relation to the merus of the second walking leg (anterior face). Large blood channels (located away from the membrane) supplying the posterior gas window (arrow) are visible behind the anterior gas-window blood system. Small blood channels (star) project towards the anterior gas-window membrane. Dorsal is top left, proximal top right. **C**, The internal surface of a complete cast. A distributing system (arrows, lower left) is linked to a collecting system (arrow, top right) via one or more 'portal' systems (stars). The blood is forced to pass from system to system through fine channels ('lacunae'; see **D**) in direct contact with the gas-window membrane. Same orientation as in **A** and **B**. **D**, The external surface of a cast showing the flat lacunar distribution of blood directly beneath the gas-window membrane. The small round holes (arrows) mark the positions of 'pillar cells' which anchor the connective tissue matrix to the membrane and which have been removed during the casting procedure.

(100–200 μm wide and 40–70 μm deep) before being redirected towards a new patch of membrane through a new set of smaller channels. The process is repeated one or more times until the blood leaves the gas-window complex via a collecting system (Fig. 2C) (compare with the crab lung portal systems⁹). In *Scopimera* most, if not all, of the venous blood returning from the leg segments distal to the merus is forced to pass through the gas-window complex (D.P.M., unpublished). The cuticle covering the gas window is extremely thin (0.6 μm) (D.P.M., unpublished) and compares well with the ultra-thin lung cuticle (0.2 μm) of other air-breathing crabs^{6,8,10}.

The total surface area of the gas window in *S. inflata* is large (Fig. 3A) and greater than the predicted lung surface area in another air-breathing crab, *Holthuisana transversa*¹¹. Some species of *Dotilla* (for example *Dotilla myctiroides*) also have gas windows on the sternal plates of the thorax³. Although *D. myctiroides* has consistently smaller leg gas windows and narrower meri than *S. inflata*, the total membrane area in *D. myctiroides* is greater because of the added area provided by the sternal gas windows (Fig. 3A).

On morphological grounds, therefore, the gas windows are well suited for aerial gas exchange. To obtain physiological evidence for such a function, the oxygen consumption of 11 male *S. inflata* was measured using Warburg respirometers¹² adapted to accommodate crabs. The gas windows were painted over with a coat of Humbrol (Hull, UK) non-toxic, oil-based enamel paint and the rate of oxygen consumption was compared with the rate before painting of the windows and with the rate after the paint had been removed (Fig. 3B). While in the flasks, the crabs were continually active, consuming large amounts of oxygen even when the flasks were kept in the dark. This response was consistent and prevented measurement of basal metabolic rates. Behavioural observations revealed that crabs always attempted to move off and away from non-sand substrates (for example, glass) or continually attempted to dig. However, when the gas windows were painted over, mean oxygen consumption ($\pm$ s.d.) fell by $59.8 \pm 8.5\%$ (range 40.9–69.9%); when the paint was removed the oxygen consumption rose significantly ($P > 0.001$, paired Student's *t*-test). Although animals sometimes did not recover to 'pre-painted' levels, the effect was nonetheless dramatic and certainly indicates that gas windows are involved in aerial gas exchange. Indeed, the fall in oxygen consumption would have been more dramatic had the crabs not been able to remove some of the paint (before it dried) by scraping their legs together—a behaviour that was difficult to prevent. When another coat of paint was applied to the legs after the first had dried, the gas windows were completely sealed and these animals always died soon after. Crabs never died, however, when the two coats of paint were perforated, and in these cases the oxygen consumption was similar to that in crabs with one coat of paint. To a certain extent, this result excludes a possible toxic effect of the paint and confirms the importance of the gas windows in aerial gas exchange.

The most characteristic structural adaptation to air-breathing in crabs has been the expansion and vascularization of the branchiostegites forming an air cavity above the gills, and thus a lung^{6,9,13}. *Scopimera* and *Dotilla* are most unusual among air-breathing crabs because they do not have a lung; the brachial chambers are small and are completely filled by the six gills which are continually bathed in water and are therefore not available for aerial gas exchange. The branchiostegites are poorly vascularized and are not expanded laterally. There is, therefore, no air cavity or vascular lining above the gills through which any significant aerial gas exchange might occur. Instead, these crabs have developed an alternative air-breathing structure—the unique, extra-branchial gas windows. These structures almost certainly serve to oxygenate venous blood returning from the leg sinuses.

Despite the obvious vulnerability of such thin and exposed membranes, gas windows have evolved in several genera of

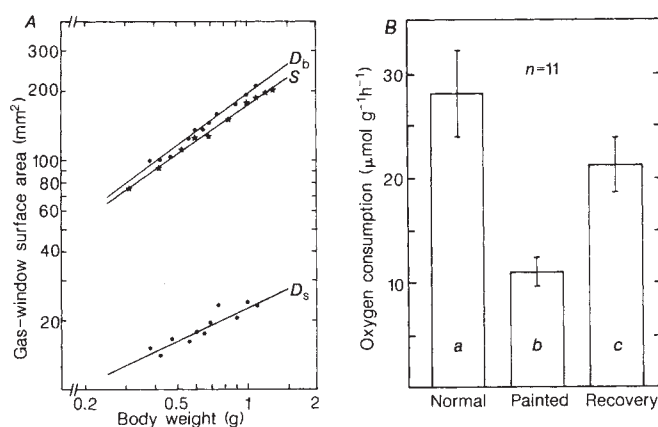


Fig. 3 **A**, Graphs showing the total surface area of gas-window membranes on legs of male *Scopimera inflata* of different body weights (S); and for *Dotilla myctiroides* the total surface area of both leg and sternal gas windows (D_b), and the area of sternal gas windows alone (D_s). Lines were fitted to the data by least-squares linear regression, $y = aM^b$ where y = the total surface area in mm², M = body wet weight (in g), a = value of y when $M = 1$ g and b = slope of regression of $\log y$ on $\log M$: for S , $y = 172.2M^{0.699}$; for D_b , $y = 190.6M^{0.730}$; for D_s , $y = 22.8M^{0.476}$. The surface area of leg gas windows alone in *D. myctiroides* follows the relationship $y = 167.8M^{0.769}$ (not shown). Surface areas were calculated by cutting out camera lucida tracings of gas-window perimeters (drawn at $\times 25$ or $\times 50$ magnification) and their weights were compared with that of a similar piece of paper of known surface area. **B**, Histograms showing the rate of oxygen consumption of 11 male *S. inflata* ((0.78–1.12 g) before painting of the gas-window membrane (a); after painting (b); and after the paint has been removed by simply peeling it off (c). Mean oxygen consumption ($\pm$ s.d.) was $27.8 \pm 4.3 \mu\text{mol g}^{-1} \text{h}^{-1}$ in a , $10.8 \pm 1.4 \mu\text{mol g}^{-1} \text{h}^{-1}$ in b , and $21.0 \pm 2.7 \mu\text{mol g}^{-1} \text{h}^{-1}$ in c ; b is significantly different from both a and c (paired t -test, $P > 0.001$). Warburg flasks of 60–62 cm³ internal volume were used. Crabs were collected daily and rates of respiration were measured in air. Crabs and apparatus were equilibrated for 30 min and experiments were run for 2 h at 28 °C. Readings were taken every 10 min and the rate of oxygen consumption (per g wet weight per h) was calculated from the mean of these values and corrected to standard conditions of temperature and pressure, dry. Normal and painted crabs were run consecutively and recovery was checked the following morning.

ocypodid crabs³. To the researcher they present a unique expanse (up to 25 mm²) of exposed, ultra-thin respiratory membrane which is readily accessible and easily manipulated. As the membranes are transparent, the mechanisms of blood flow through a capillary-like network can also be observed directly. Thus, they represent a useful experimental system.

I thank the Australian Institute of Marine Science, Townsville, for the use of facilities during oxygen consumption measurements; P. Ng and L. Tan at The Department of Zoology, National University of Singapore for their generous hospitality; J. Zeil, C. Farrelly, D. Sandeman and P. Greenaway for useful discussion; and R. Foertmeyer for translations from the German literature. This work was supported by an ARGS grant to D. C. Sandeman.

Received 16 September; accepted 15 November 1985.

- McCulloch, A. R. & McNeill, F. A. *Rec. Aust. Mus.* **14**, 49–59 (1923–1926).
- Hartnoll, R. G. *Estuar. coast. mar. Sci.* **1**, 137–152 (1973).
- Kemp, S. *Rec. Indian Mus.* **16**, 305–348 (1919).
- Aurivillius, C. W. S. *Nova Acta R. Soc. Scient. Upsal. Ser. 3* **16**, 1–43 (1893).
- de Haan, W. in *Fauna Japonica* (ed. de Siebold, Ph. Fr.) 24–25 (J. Muller & Sons, Amsterdam, 1833–1850).
- Greenaway, P. & Farrelly, C. A. *J. Morph.* **181**, 133–142 (1984).
- Maynard, D. M. *The Physiology of Crustacea* Vol. 1, 161–226 (Academic, London, 1960).
- Storch, V. & Welsch, U. *Mar. Biol. (Berl.)* **29**, 363–371 (1975).
- von Raben, K. *Z. wiss. Zool.* **145**, 425–461 (1934).
- Taylor, H. H. & Greenaway, P. *J. Zool.* **189**, 359–384 (1979).
- Greenaway, P. *Aust. J. Zool.* **32**, 1–6 (1984).
- Umbreit, W. W., Burris, R. H. & Stauffer, J. F. *Manometric and Biochemical Techniques* 5th edn (Burgess, Minnesota, 1972).
- Diaz, H. & Rodriguez, G. *Biol. Bull. mar. biol. Lab., Woods Hole* **153**, 485–504 (1977).

Ganglion cell dendrites are presynaptic in catfish retina

Hiroko M. Sakai, Ken-Ichi Naka & John E. Dowling

National Institute for Basic Biology, Okazaki 444, Japan

The retinal ganglion cells are third-order, spike-generating neurones whose axons transmit the output of the retina to the rest of the brain. It has long been believed that the dendrites of the retinal ganglion cells, like the dendrites of most other Golgi type I neurones, are only postsynaptic. Here we have studied the synapses made onto the ganglion cells in the catfish (*Ictalurus punctatus*), and we report that the distal dendrites of large-field ganglion cells make conventional chemical synapses onto other inner plexiform layer processes. We have also found that, more than 100 µm away from the cell perikaryon, the synapses made onto and by these ganglion cell dendrites are often tightly clustered. These synaptic clusters appear to be quite regularly spaced along the dendrites. Our results have important implications for the identification of ganglion cell dendrites within the inner plexiform layer as well as for the understanding of the ganglion cell response and receptive field generation^{1,2}.

Ganglion cells in the catfish eyecup preparation were filled with horseradish peroxidase (HRP) either by intracellular injection following the recording of their physiological responses, or by back-filling by immersion of the cut optic nerve in Ringer's containing HRP for 3–5 h. Then, the eyecup was fixed in 4% paraformaldehyde and 1.25% glutaraldehyde in 0.1 M cacodylate buffer overnight. The retina was removed from the eyecup, reacted with diaminobenzidine, postfixed with 2% osmium tetroxide and embedded in Epon. The retina was sectioned tangentially, initially at a thickness of 30 µm, and stained ganglion cells were photographed and drawn. Sections containing one intracellularly stained ganglion cell, or several back-filled ganglion cells, were glued to blank Epon blocks, and thin (800 Å) serial sections were cut, mounted on Formvar-coated slot grids and stained with uranyl acetate and lead citrate for electron microscopy³.

Figure 1A shows a photograph and a camera lucida drawing of a large-field ganglion cell injected with HRP. This cell fired spikes at both the onset and cessation of a step of light, that is, it was an 'on-off' ganglion cell^{4,5}. Its perikaryon was located in the ganglion cell layer and its dendritic tree was bistratified, with processes running in both sublamina a and b of the inner plexiform layer⁶. The dendritic field of the cell was over 1 mm in diameter. An axon-like process was seen emerging from the perikaryon and running towards the optic nerve (not shown). The cell perikaryon and dendrites were examined by electron microscopy for evidence of synaptic junctions. The results (Fig. 1A) were typical of the large-field ganglion cells examined so far. The perikaryon and proximal dendrites were only postsynaptic. For the cell shown in Fig. 1A, all but one of the synapses made on the proximal dendrites were of the conventional type (solid arrowheads), indicating that these were likely to be amacrine cell synapses¹. One synaptic junction was of the ribbon type (open arrowhead), indicating that it was a bipolar cell junction¹. In other cells, a more even mix of bipolar and amacrine cell input was observed on the proximal dendrites, and synapses were also found on the cell perikaryon.

Beyond about 100 µm from the cell perikaryon, clusters of synaptic contacts were observed along the dendrites in sublamina a and b . These clusters consisted of 3–7 synaptic contacts spread along 10–30 µm of dendrite. Along one dendrite, which was followed for over 400 µm, three clusters were observed that were each separated by ~100 µm. We observed both amacrine (conventional) and bipolar (ribbon) synapses onto the ganglion cell dendrites in every cluster. In addition, synapses made by

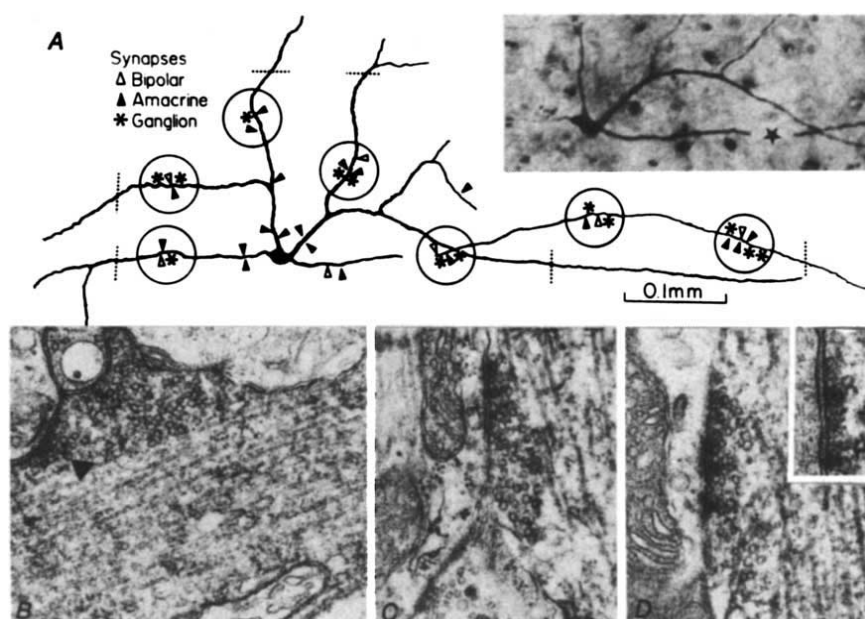


Fig. 1 A, Camera lucida drawing of an HRP-stained, large-field, on-off ganglion cell contained within a 30-μm tangential section of the catfish retina. One series of thin sections encompassed the entire cell perikaryon and the dendrites out to the dotted lines. Synapses made onto and by the cell are indicated approximately where they were observed on the cell. Each circled area represents a cluster of input and output synapses. Inset, a light micrograph showing part of the cell. The star indicates a small portion of the dendrite that was not included in the 30-μm-thick section of the cell. B, C, Electron micrographs showing synapses made by the ganglion cell in A. In B the synapse is marked by an arrowhead. D, Synapses observed in back-filled ganglion cell dendrites. Magnifications: B, $\times 23,000$; C, $\times 21,000$; D, $\times 26,000$; inset, $\times 32,000$.

the ganglion cell dendrites were observed at these sites. As such synaptic junctions have not been seen previously, they will be described in detail.

Immediately adjacent to the synapses made by the ganglion cell dendrites, a dense aggregation of synaptic vesicles was typically observed. Often the dendrite bulged outwards slightly at such places (Fig. 1B), but this was not always the case (Fig. 1C). At the presumed site of synaptic contact, electron-dense material was usually observed in association with the membranes of the contacting elements and within the synaptic cleft (Fig. 1B–D). Synaptic vesicles were never observed in the ganglion cell dendrites away from the synaptic sites. The synapses made by the ganglion cell dendrites were often on processes that contained no synaptic vesicles. However, we have observed ganglion cell synapses on vesicle-containing (amacrine or ganglion cell) processes and on ribbon-containing (bipolar cell) processes. As yet, a ganglion cell synapse on a process that was itself synapsing on the dendrite (a reciprocal synapse) has not been observed. It was not possible to observe synapses made by the dendrites in every large-field HRP-stained ganglion cell. If the HRP stain was very dense, intracellular structures were obscured and it was impossible to observe the synapses. We have been successful, however, in identifying dendritic synapses in those large-field cells that were stained dark enough to be positively identified in the electron microscope, but light enough for the cell interior to be visualized. So far, we have observed synapses made by the distal dendrites of the following three large-field ganglion cells injected intracellularly with HRP: the bistratified 'on-off' cell shown in Fig. 1A; a monostратified, large-field, on-centre ganglion cell; and a monostратified, large-field, off-centre ganglion cell⁷.

Figure 2 shows data recorded from the large-field, off-centre ganglion cell; this cell was spontaneously active and responded to full-field illumination with a sustained hyperpolarization followed by transient depolarizations on removal of the light (Fig. 2A). The cell perikaryon was located in the ganglion cell layer, and its dendrites were confined to sublamina *b* of the inner plexiform layer. The dendritic field of this cell also extended over 1 mm in diameter. Conventional (amacrine cell) synapses were seen on the cell perikaryon, while on the proximal dendrites both conventional and ribbon (bipolar cell) synapses were found. Beyond about 100 μm from the cell perikaryon, synapses made by the ganglion cell dendrites were observed and, as was the case for the on-off cell shown in Fig. 1, clusters

of synapses were observed along the distal dendrites. Figure 2C shows an example of a synapse made by this cell which was found about 150 μm from the perikaryon (arrow in Fig. 2B).

Although each of the large-field ganglion cells studied so far gave typical ganglion cell responses, with spikes of amplitude 5–20 mV, an axon was positively identified only for the cell shown in Fig. 1A. To ensure that we had correctly identified these neurones as ganglion cells, we examined three series of sections containing back-filled ganglion cells. In each of these series, synapses made by stained dendrites, similar to those observed in the intracellularly stained cells, were seen. Figure 1D shows two such examples. The preservation of the back-filled ganglion cells was not usually as satisfactory as that of the intracellularly injected cells; nevertheless, the results were unequivocal in showing synaptic junctions made by the dendrites of back-filled ganglion cells.

We have also studied two small-field ganglion cells, filled intracellularly with HRP, whose dendritic fields were, respectively, ~200 and 350 μm in diameter. Synapses onto the perikarya and dendrites of these cells were readily observed, but the synapses did not appear to be clustered along the dendrites as was the case for the large-field cells. Furthermore, no synapses made by the dendrites of these ganglion cells have been observed. The HRP staining of these cells was denser than that of the large-field ganglion cells described above, and intracellular structures in some of the dendrites were obscured. However, about two-thirds of the dendrites were stained lightly enough for the synapses to have been observed if they were present. Thus, it seems likely that synapses are made only by the distal dendrites of the large-field ganglion cells in catfish.

It has long been recognized that the dendrites of many short-axon or axonless (Golgi type II) cells in the retina and elsewhere may be both pre- and postsynaptic^{1,8–10}. Many of these cells have been shown to function mainly with sustained, graded potentials^{11,12}. Only in a few instances have presynaptic dendrites been observed in long-axon cells^{8,13}, and there has been no evidence so far that any of the long-axon ganglion cells of the retina, which are typical Golgi type I cells, have dendrites which are presynaptic. An obvious and important question to ask is whether ganglion cells in other species, and other Golgi type I cells, make dendritic synapses. It is of interest to note in this regard that there is recent physiological evidence for direct interactions between ganglion cells in the cat retina¹⁴. These authors argued, however, that these interactions are more likely

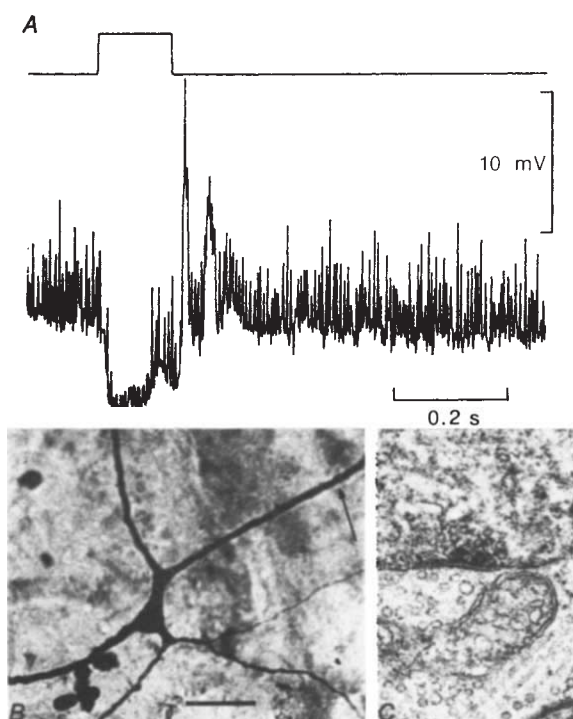


Fig. 2 A, Step-evoked response from a large-field, off-centre ganglion cell. The cell showed spontaneous activity which was depressed during the light-evoked hyperpolarization. The transient depolarizations that occurred during the sustained hyperpolarization are characteristic of off-centre ganglion cells in catfish. Two large sharp depolarizations occurred on removal of the light. The smaller amplitude of the spikes compared with the excitatory postsynaptic potentials seen in this cell suggests that the spikes must have been produced some distance away from the recording site. The light stimulus was $0.1 \mu\text{W cm}^{-2}$. B, Light micrograph showing part of the large-field off-centre ganglion cell whose response to light is shown in A. Scale bar, $50 \mu\text{m}$. C, An electron micrograph showing a synapse made by a dendrite of the ganglion cell shown in B. The site of the synapse is indicated by an arrow in B. $\times 26,000$.

to be mediated by electrical junctions than by chemical synapses.

The clustering of synapses along the distal dendrites of the catfish large-field ganglion cells is a novel observation which has implications for the generation of ganglion cell responses and for inner plexiform layer information processing. Similar clustering of pre- and postsynaptic sites has been observed in certain leech and insect neurones¹⁵⁻¹⁷. That the clusters are separated by $100 \mu\text{m}$ or more along the ganglion cell dendrites suggests that these synaptic aggregations operate relatively independently of one another and may represent sites for local signal processing within the inner plexiform layer².

As noted above, the dendritic fields of many of the catfish ganglion cells extend well over 1 mm in diameter. It is unlikely that single synaptic inputs several hundred micrometres away from the cell perikaryon will significantly affect the spike output of the cell. Thus, the clusters of synapses observed along the distal dendrites may also serve as integrating centres such that, when acting in concert in a cluster, the distal synapses are capable of shaping the ganglion cell responses.

J.E.D. was on leave of absence from Harvard University and was supported by a grant from the Japanese Society for the Promotion of Science.

4. Hartline, H. K. *Am. J. Physiol.* **121**, 400-415 (1938).
5. Chan, R. Y. & Naka, K.-I. *Vision Res.* **16**, 1119-1129 (1976).
6. Nelson, R., Famiglietti, E. V. Jr & Kolb, H. *J. Neurophysiol.* **41**, 472-483 (1978).
7. Davis, G. W. & Naka, K.-I. *J. Neurophysiol.* **43**, 807-831 (1980).
8. Rall, W., Shepherd, G. M., Rees, T. S. & Brightman, M. W. *Expl. Neurol.* **14**, 44-56 (1966).
9. Ralston, H. J. III *Nature* **230**, 585-587 (1971).
10. Ralston, H. J. III in *The Neurosciences: 4th Study Program* (eds Schmitt, F. O. & Worden, F. G.) 373-379 (MIT Press, 1979).
11. Werblin, F. S. & Dowling, J. E. *J. Neurophysiol.* **32**, 339-355 (1969).
12. Kaneko, A. *J. Physiol., Lond.* **207**, 623-633 (1970).
13. Hamos, J., Van Horn, S. C., Baczkowski, D., Uhlrich, D. J. & Sherman, S. M. *Nature* **317**, 618-621 (1985).
14. Mastrorade, D. W. *J. Neurophysiol.* **49**, 350-365 (1983).
15. Muller, K. J. & McMahan, U. J. *Proc. R. Soc. B194*, 481-499 (1976).
16. Watson, A. H. D. & Burrows, M. *J. comp. Neurol.* **205**, 383-397 (1982).
17. Watson, A. H. D. *J. Neurocytol.* **13**, 303-327 (1984).

Different factors from the central nervous system and periphery regulate the survival of sensory neurones

Alun M. Davies, Hans Thoenen & Yves-Alain Barde

Department of Anatomy, St George's Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK and Department of Neurochemistry, Max Planck Institute for Psychiatry, 8033 Planegg-Martinsried, München, FRG.

Work on nerve growth factor has established that the survival of developing vertebrate neurones depends on the supply of a neurotrophic factor from their target field¹⁻⁵. The discovery of several new neurotrophic factors⁶ has raised the possibility that neurones which innervate multiple target fields require several different neurotrophic factors for survival. Here we show that two distinct neurotrophic factors, one in the central nervous system (CNS) and the other in skeletal muscle, promote the survival of proprioceptive neurones in culture. At saturating concentrations, either factor alone supported most neurones and there was no additional survival in the presence of both factors, but at sub-saturating concentrations the combined effect was additive. The neurotrophic activity of each factor was greatest during the period of natural neuronal death. Our results demonstrate that each cultured proprioceptive neurone responds to two distinct neurotrophic factors present in its respective central and peripheral target fields, and suggest that these factors cooperate in regulating survival during development.

To investigate the specific neurotrophic requirements of neurones that innervate two distinct target fields, we studied the neurones of the embryonic chick trigeminal mesencephalic nucleus (TMN). Despite their location in the brain stem, these are unipolar, first-order, sensory neurones derived from the neural crest⁷, with processes that divide into central and peripheral branches. The central branches innervate certain brainstem nuclei, including the trigeminal motor nucleus, and the peripheral branches are distributed to proprioceptors in the muscles of mastication⁸. *In vivo*, 70% of TMN neurones die during the innervation of these central and peripheral target fields⁹.

We used a recently described method to establish glia-free cultures of TMN neurones¹⁰. The survival and growth of most of the neurones was promoted by an extract of skeletal muscle, their specific peripheral target tissue (Fig. 1). Because these neurones also project to target fields in the CNS, we sought to ascertain whether a recently purified protein¹¹, brain-derived neurotrophic factor (BDNF), known to act on nodose and dorsal root ganglion neurones¹², also promotes the survival of TMN neurones. As with skeletal muscle extract, we found that most TMN neurones (70-80%) survived and grew in the presence of saturating concentrations of BDNF (Fig. 1) and could be sustained in culture for more than a month with regular supplements of either factor. Virtually no neurones survived without BDNF or skeletal muscle extract. The combination of BDNF and

Received 16 September; accepted 21 November 1985.

1. Dowling, J. E. & Boycott, B. B. *Proc. R. Soc. B166*, 80-111 (1966).
2. Koch, C., Poggio, T. & Torre, V. *Phil. Trans. R. Soc. B298*, 227-264 (1982).
3. Sakai, H. & Naka, K.-I. *Vision Res.* **23**, 339-351 (1983).

muscle extract was additive at concentrations below saturation (Fig. 1), whereas at saturating concentrations, the combination did not promote the survival of additional neurones, indicating that all the surviving neurones responded to both factors. As the neurotrophic factor in muscle extract has not yet been purified it was necessary to show that it is functionally distinct from BDNF. This was done using explants of the ventrolateral part of the trigeminal ganglion. Although BDNF elicited profuse outgrowth from these explants, skeletal muscle extract did not (Fig. 2). In contrast, both agents were active on TMN explants (Fig. 2).

The neurite-promoting effects of BDNF and muscle extract on TMN explants showed a very similar developmental time-

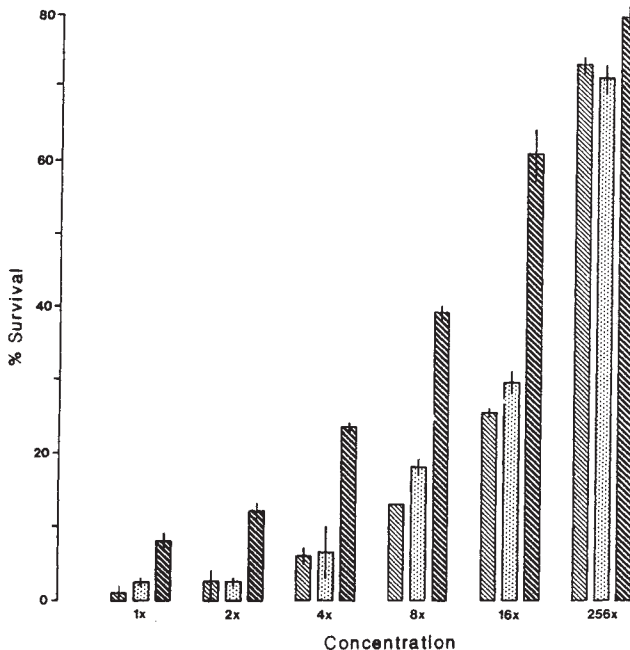


Fig. 1 Dose-responses of TMN neurones to BDNF (lightly hatched bars), skeletal muscle extract (stippled bars) and BDNF + muscle extract (heavily hatched bars). BDNF, 1x = 25 $\mu\text{g ml}^{-1}$; muscle extract, 1x = 625 ng ml^{-1} ; BDNF + extract, 1x = 25 $\mu\text{g ml}^{-1}$ + 625 ng ml^{-1} . The median region of the tectal commissure of day 10 (E10) chick embryos was trypsinized and dissociated by gentle trituration and TMN neurones were purified by differential sedimentation at 1g as described previously¹⁰. The yield of glia-free TMN neurones by this method was >60%. We plated 1,000–2,000 neurones in 35-mm culture dishes containing 2 ml F14 medium¹² with 10% heat-inactivated horse serum. Before culture the dishes were coated with polyornithine and then laminin¹⁸. BDNF was purified as described elsewhere¹¹ using the modifications described by Lindsay *et al.*¹². Skeletal muscle extract was prepared from E18 chick pectoral muscle: 100 g muscle was homogenized in 250 ml 10 mM phosphate buffer at pH 7.0 containing 150 mM NaCl and 0.5 mM EDTA using a Sorvall Omnimixer. This homogenate was centrifuged at 20,000g for 1 h and ammonium sulphate added to the supernatant to 30% saturation. The resulting precipitate was discarded and ammonium sulphate increased to 50% saturation. This precipitate was dissolved in water and passed through a Sephadex G25 column equilibrated with F14 medium from which the void fraction containing the protein was collected. The cultures were viewed by phase-contrast microscopy after 48 h incubation and the percentage of neurones with processes >10 cell diameters was determined. The results are expressed in percentage survival, 100% being the number of cells at the time of plating. 2% of the neurones survived in the absence of BDNF and muscle extract (control cultures), and this value is subtracted from all bars. The mean and range of the percentage survival at each concentration is plotted, $n=2$. NGF (0.2–625 ng ml^{-1}), either alone or in combination with BDNF, did not promote or influence neuronal survival (A.M.D., A. G. S. Lumsden and H. Rohrer, unpublished data).

course (Fig. 3), which suggests that these factors are active during the same stage of neuronal development. The magnitude of neurite outgrowth in response to each factor was most marked at stages E10 and E11, which are closely related to the stage of natural neuronal death in the TMN (E9–E13)¹³. Thus, TMN neurones are most sensitive to the neurotrophic effects of BDNF and muscle extract at the same time as their survival is critically dependent on forming appropriate central and peripheral connections.

Our findings suggest that the survival of developing sensory neurones *in vivo* is normally dependent on the supply of both a neurotrophic factor from the periphery and a different

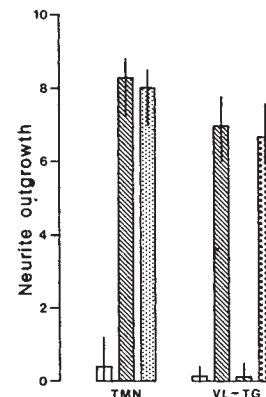


Fig. 2 Magnitude of neurite outgrowth from explants of the ventrolateral part of the trigeminal ganglion (VL-TG) and from TMN explants in control cultures (clear bars) and cultures supplemented with BDNF at 100 ng ml^{-1} (hatched bars), skeletal muscle extract at 160 $\mu\text{g ml}^{-1}$ (light stippled bars) and BDNF + muscle extract (heavy stippled bars). Five explants were embedded in 250 μl collagen gel in 16-mm diameter plastic wells¹⁹ and overlaid with 100 μl heat-inactivated horse serum and 650 μl F14 medium. The cultures were examined after 24 h incubation and the magnitude of neurite outgrowth was scored on an ordinal scale²⁰ from 0 to 10. The median and interquartile deviation of these scores are plotted. Although BDNF elicited marked fibre outgrowth from both VL-TG and TMN explants, skeletal muscle extract was active only on TMN explants. Note that BDNF was also active on VL-TG explants in the presence of muscle extract, showing that the lack of response of VL-TG explants to muscle extract alone is not caused by an inhibitory effect.

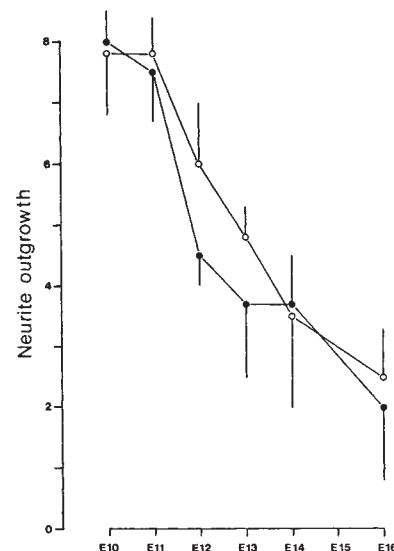


Fig. 3 Magnitude of neurite outgrowth from TMN explants of E10–E16 chick embryos after 24 h incubation with either BDNF at 100 ng ml^{-1} (solid circles) or muscle extract at 160 $\mu\text{g ml}^{-1}$ (open circles). The median and interquartile deviation are shown for each age.

neurotrophic factor from the CNS. Note that Yip and Johnson^{13,14} have demonstrated that peripheral or central rhizotomy of dorsal root ganglia in neonatal rats results in the death of about half of the neurones. Furthermore, they showed that NGF administered at the time of the lesion prevented neurone death. In view of our finding that sensory neurones can be sustained by a sufficiently high concentration of either a peripheral or a central neurotrophic factor alone (Fig. 1), we suggest that NGF (found *in vivo* in the peripheral but not the central branches of dorsal root ganglia¹⁵) administered in high enough concentrations may rescue all NGF-responsive sensory neurones even when experimentally deprived of the supply of their CNS-derived trophic factor which we believe to be either BDNF or another, as yet uncharacterized, molecule.

Our demonstration of the response of sensory neurones to distinct central and peripheral neurotrophic factors provides the basis of a mechanism for selectively supporting the survival of only those neurones which make appropriate connections in both target fields. If the availability of each neurotrophic factor in the respective central and peripheral target fields lies below the critical threshold for maintaining survival, it would be necessary for a developing sensory neurone to make both the appropriate central and peripheral connections to procure sufficient trophic support to ensure survival. Supporting this contention is evidence from immunoassays of NGF¹⁶ and from the purification of BDNF¹¹ that neurotrophic factors are normally present in only minute amounts. Furthermore, the amount of NGF present in target organs does not allow the survival of all innervating neurones because additional NGF rescues neurones that would otherwise die during development¹⁷. Addi-

tional NGF also increases the levels of catecholamine-synthesizing enzymes², indicating that the synthesis of these enzymes is not normally fully stimulated by NGF *in vivo*. Also, there is a correlation between the density of sympathetic innervation and the tissue concentration of NGF¹⁶ and NGF messenger RNA^{4,5}, suggesting that the density of innervation is regulated by the availability of neurotrophic factor.

In the light of our findings it will be of interest to ascertain whether, in addition to sensory neurones, other neurones which innervate more than one target field require a different neurotrophic factor from each field for survival.

We thank the Wellcome Trust for the award of a Research Travel grant to A.M.D. and Frau Čáp for technical assistance.

Received 25 July; accepted 10 December 1985.

1. Levi-Montalcini, R. & Angeletti, P. U. *Physiol. Rev.* **48**, 534-569 (1968).
2. Thoenen, H. & Barde, Y.-A. *Physiol. Rev.* **60**, 1284-1335 (1980).
3. Cowan, W. M., Fawcett, J. W., O'Leary, D. D. M. & Stanfield, B. B. *Science* **225**, 1258-1265 (1984).
4. Heumann, R., Korsching, S., Scott, J. & Thoenen, H. *EMBO J.* **3**, 3183-3189 (1984).
5. Shelton, D. L. & Reichardt, L. F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7951-7955 (1984).
6. Barde, Y.-A., Edgar, D. & Thoenen, H. *A. Rev. Physiol.* **45**, 601-612 (1983).
7. Narayanan, C. H. & Narayanan, Y. *J. embryol. exp. Morph.* **43**, 85-105 (1981).
8. Manni, E., Bortolani, R. & Azzena, G. B. *Expl. Neurol.* **12**, 320-328 (1965).
9. Rogers, L. A. & Cowan, W. M. *J. comp. Neurol.* **147**, 291-320 (1973).
10. Davies, A. M. *Dev. Biol.* (in the press).
11. Barde, Y.-A., Edgar, D. & Thoenen, H. *EMBO J.* **1**, 549-553 (1982).
12. Lindsay, R. M., Thoenen, H. & Barde, Y.-A. *Dev. Biol.* **112**, 319-328 (1985).
13. Yip, K. H. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6245-6249 (1984).
14. Johnson, E. M. & Yip, K. H. *Nature* **314**, 751-752 (1985).
15. Korsching, S. & Thoenen, H. *Neurosci. Lett.* **54**, 201-205 (1985).
16. Korsching, S. & Thoenen, H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3513-3516 (1983).
17. Hamburger, V., Brunso-Bechtold, J. K. & Yip, J. W. *J. Neurosci.* **1**, 60-71 (1981).
18. Edgar, D., Timpl, R. & Thoenen, H. *EMBO J.* **3**, 1463-1468 (1984).
19. Lumsden, A. G. S. & Davies, A. M. *Nature* **306**, 786-788 (1983).
20. Davies, A. M. & Lumsden, A. G. S. *J. comp. Neurol.* **223**, 124-137 (1984).

Proliferating bipotential glial progenitor cells in adult rat optic nerve

Charles French-Constant & Martin C. Raff

Medical Research Council Neuroimmunology Project,
Zoology Department, University College London,
London WC1E 6BT, UK

We have shown previously that the rat optic nerve contains three types of macroglial cells—oligodendrocytes and two types of astrocytes—which develop as two distinct lineages. Type-1 astrocytes develop from one type of precursor cell beginning at embryonic day 16 (E16)¹, while oligodendrocytes and then type-2 astrocytes develop from a common, bipotential progenitor cell beginning at birth (E21) and postnatal days 7-10 (P7-10), respectively². Here we report that proliferating bipotential oligodendrocyte-type-2 astrocyte (0-2A) progenitor cells are present in the adult rat optic nerve, raising the possibility that these cells are produced continually from self-renewing stem cells throughout life.

0-2A progenitor cells in cultures of perinatal optic nerve have a characteristic process-bearing morphology³⁻⁶ and are labelled by the A2B5 monoclonal antibody, which recognizes specific gangliosides⁷, but not by antisera against glial fibrillary acidic protein (GFAP), which react specifically with astrocytes^{8,9}, or by monoclonal anti-galactocerebroside (GC) antibody¹⁰, which reacts specifically with oligodendrocytes^{10,11}. Although 0-2A progenitor cells in the perinatal optic nerve proliferate *in vivo*¹², *in vitro* they prematurely stop dividing and differentiate. In cultures containing $\geq 10\%$ fetal calf serum (FCS), most acquire GFAP within 1-2 days to become A2B5⁺, GFAP⁺ type-2 astrocytes, whereas in cultures containing $\leq 1\%$ FCS most acquire GC within 2 days to become A2B5⁺, GC⁺ oligodendrocytes^{13,14}, which soon become A2B5⁻ (ref. 5). On the other hand, when cultured on a monolayer of type-1 astrocytes in 0.5% FCS, or in low-serum medium conditioned by type-1

astrocytes, 0-2A progenitor cells continue to proliferate and produce post-mitotic oligodendrocytes for weeks^{4,5}, just as they do *in vivo*^{15,16}. Moreover, when E17 optic nerve cells are cultured on type-1 astrocytes in 0.5% FCS, the first oligodendrocytes appear after 4 days *in vitro*⁵, equivalent to the time they first appear *in vivo*². These observations led us to propose that the timing of oligodendrocyte differentiation depends on an intrinsic clock in the 0-2A progenitor cells. The clock operates by counting cell divisions, which are driven by growth factors released by type-1 astrocytes. After undergoing a certain number of divisions, the cells become unresponsive to the growth factors, stop dividing and, as a consequence, differentiate into oligodendrocytes. Type-2 astrocyte differentiation, on the other hand, we propose requires an inducing factor (which can be mimicked by 10% FCS in culture), which may not appear in the developing optic nerve until after P7⁵.

The present study was undertaken to determine whether 0-2A progenitor cells are present in the adult rat optic nerve. Cells from optic nerves of adult Sprague-Dawley rats (4-6 months old, 300-350 g) were dissociated and cultured on type-1 astrocyte monolayers in low-serum conditions (<1% FCS) as described in Table 1 legend. Staining with A2B5 antibody after 2-4 days in culture revealed A2B5⁺ cells showing a characteristic bipolar or tripolar 0-2A progenitor-cell-like morphology (Fig. 1); in double-labelling experiments these cells were GFAP⁻ (Fig. 1) and GC⁻ (not shown). Such cells were also seen in cultures of optic nerves from rats that were more than 1 year old. Although most GFAP⁺ astrocytes in adult optic nerve are A2B5⁺, very few cells with this antigenic phenotype were seen in our experiments, presumably because they were not released from the nerves by our dissociation procedure¹⁷.

The A2B5⁺ cells showing 0-2A progenitor-cell-like morphology usually occurred in groups of two or more (Fig. 1) and their numbers increased with time (doubling about once a day), suggesting that they were proliferating (Table 1). This was confirmed by autoradiographic studies, in which more than 60% of these cells were radiolabelled following an 18-h pulse of ³H-thymidine (Table 1). Similar results were obtained when

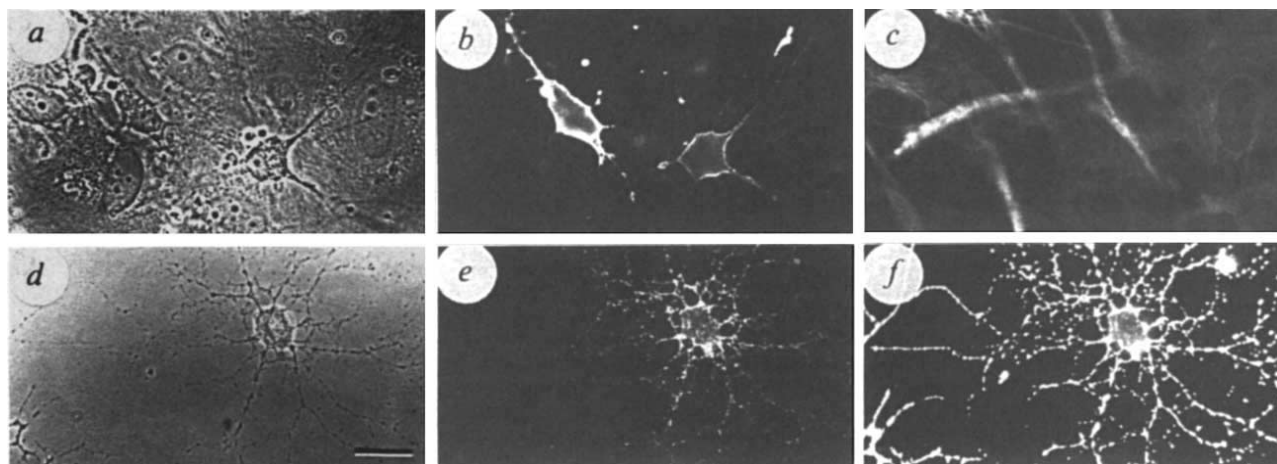


Fig. 1 *a-c*, Two A2B5⁺, GFAP⁻ 0-2A progenitor-like cells from adult optic nerve growing on a monolayer of type-1 astrocytes. The cells were photographed using phase contrast (*a*), rhodamine (*b*) and fluorescein (*c*) optics. Note that while the 0-2A progenitor-like cells are A2B5⁺, GFAP⁻, the underlying astrocytes are A2B5⁻, GFAP⁺. *d-f*, An A2B5⁺, GC⁺ oligodendrocyte in an adult optic nerve culture switched from ACM to B-S/0.5% FCS for 3 days. Note that the processes of an adjacent oligodendrocyte on the left of the photograph (whose cell body is out of the field) are GC⁺ but A2B5⁻. Scale bar, 20 μ m.

Methods. *a-c*, The cells were prefixed in 4% paraformaldehyde in phosphate buffer and then double-labelled with A2B5 antibody⁷ (culture supernatant diluted 1:2) followed by rhodamine-conjugated goat anti-mouse immunoglobulin (Cappel, diluted 1:100) and then, after fixation in acid-alcohol, with anti-GFAP serum²⁴ (1:1,000) followed by fluorescein-conjugated sheep anti-rabbit immunoglobulin (Wellcome, diluted 1:100). Finally the coverslips were mounted in glycerol containing 1,4-diazo-bicyclo(2,2,2)octane to inhibit fading of the fluorescence²⁵, examined and photographed with Tri-X film rated 400 ASA, as described previously⁹. *d-f*, Adult optic nerve cells were cultured on a PLL-coated coverslip for 3 days in astrocyte-conditioned medium and then for 3 days in B-S/0.5% FCS; the cells were labelled with monoclonal anti-GC antibody¹⁰ (ascites fluid diluted 1:100) followed by immunoglobulin-class-specific fluorescein-conjugated goat anti-IgG3 (Nordic, diluted 1:100) and then A2B5 antibody followed by immunoglobulin-class-specific rhodamine-conjugated goat anti-IgM (Cappel, diluted 1:25), as described previously⁵. Coverslips for cell counting were prefixed with paraformaldehyde to minimize cell loss during antibody labelling. The cells were mounted and photographed as described above.

cultures were maintained in low-serum medium conditioned by type-1 astrocytes (ACM) (Table 1). On the other hand, when adult optic nerve cells were cultured directly on coverslips in low-serum medium alone, the number of A2B5⁺ process-bearing cells did not increase with time, and less than 5% of such cells were radiolabelled by ³H-thymidine (Table 1). These findings suggest that 0-2A progenitor-like cells in adult optic nerve are stimulated to proliferate by growth factors released from type-1 astrocytes, as previously shown for 0-2A progenitor cells in perinatal optic nerve^{4,5}.

To determine whether the A2B5⁺, GFAP⁻, GC⁻ process-bearing cells in adult optic nerve cultures were 0-2A progenitor cells, adult nerve cells were cultured in ACM for 3 days followed by a further 3 days in non-conditioned medium containing either 0.5% or 10% FCS, then double-labelled with A2B5 and anti-GFAP antibodies or with A2B5 and anti-GC antibodies. After the initial period in ACM, more than 75% of the antibody-labelled cells were A2B5⁺, GFAP⁻, GC⁻ (Table 2) and showed an 0-2A progenitor-cell-like morphology. (In contrast to the experiments in which cells were plated directly onto type-1 astrocytes, ~20% of the A2B5⁺ process-bearing cells were weakly GFAP⁺ (Table 2), suggesting that some 0-2A progenitor cells acquired GFAP in these culture conditions.) After switching into medium containing 0.5% FCS, more than 75% of the antibody-labelled cells were GC⁺ and had a characteristic, multipolar, oligodendrocyte-like morphology. Few A2B5⁺, GFAP⁻, GC⁻ cells remained, suggesting that most such cells had acquired GC to become oligodendrocytes (Table 2). This interpretation was supported by the finding that 30–40% of the GC⁺ cells were A2B5⁺ (Fig. 1); because oligodendrocytes rapidly become A2B5⁻ as they mature^{5,18}, this finding suggests that these A2B5⁺, GC⁺ cells had only recently acquired GC and were newly developed oligodendrocytes. On the other hand, when cultures were placed in medium containing 10% FCS, more than 75% of the antibody-labelled cells became A2B5⁺, GFAP⁺ type-2 astrocytes; again few, if any, A2B5⁺, GFAP⁻, GC⁻ cells remained (Table 2). Although some of the type-2 astrocytes seen in the latter cultures may have been derived from pre-existing type-2

astrocytes rather than 0-2A progenitor cells, it seems likely that most developed from 0-2A progenitor cells that acquired GFAP in the presence of 10% FCS.

The simplest interpretation of these results is that the adult rat optic nerve contains 0-2A progenitor cells that can differentiate into either oligodendrocytes or astrocytes depending on the culture conditions. Both in this respect and in their proliferative response *in vitro* to growth factors released by type-1 astrocytes, these progenitor cells resemble those in perinatal optic nerve. However, there may be important differences between these cells in adult and perinatal nerve. Wolswijk and Noble independently have found bipotential 0-2A progenitor-like cells in 1-yr-old rat optic nerve which proliferate in response to type-1-astrocyte-derived growth factors and find that only a minority of these cells are labelled by anti-vimentin antibody (G. Wolswijk and M. Noble, in preparation). This is in contrast to our earlier findings in perinatal nerve in which the majority of 0-2A progenitor cells were vimentin-positive, and appeared to lose vimentin only after they had differentiated into oligodendrocytes¹⁴.

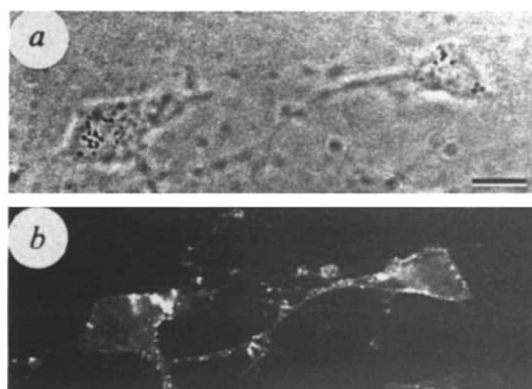
To determine whether 0-2A progenitor cells proliferate in adult animals, we injected ³H-thymidine into adult rats 18 h and 1 h before killing them and culturing cells from their optic nerves on type-1 astrocyte monolayers for 2–3 days. Between 5 and 7% (mean 6% $\pm$ 1.2 in 3 separate animals) of the 0-2A progenitor cells were radiolabelled in such experiments (Fig. 2). Although radiolabelled 0-2A progenitor cells may have divided more often in culture than unlabelled progenitor cells, thereby increasing the labelling index, these results indicate that at least some 0-2A progenitor cells proliferate in the adult optic nerve.

What is the source of proliferating 0-2A progenitor cells in the adult rat optic nerve? It seems unlikely that the progeny of individual 0-2A progenitor cells continue to proliferate for months after developing from a pre-0-2A progenitor cell in the perinatal optic nerve, since *in vitro* experiments suggest that perinatal 0-2A progenitor cells undergo few divisions before they stop proliferating and differentiate. For example, when single 0-2A progenitor cells were cultured on type-1 astrocyte monolayers in 0.5% FCS in microwells, they divided a maximum

Table 1 Effect of type-1 astrocytes on the proliferation of, and ^3H -thymidine incorporation by, 0-2A progenitor-like cells in culture

Culture substrate	No. of A2B5 ⁺ process-bearing cells after 2 or 4 days in culture		% Of A2B5 ⁺ process-bearing cells radiolabelled by ^3H -thymidine
	2 days	4 days	
Poly-L-lysine	65 $\pm$ 9	62 $\pm$ 7	1 $\pm$ 1
Type-1 astrocyte monolayer	79 $\pm$ 26	336 $\pm$ 94	62 $\pm$ 3
Poly-L-lysine + astrocyte conditioned medium	178 $\pm$ 15	302 $\pm$ 19	40 $\pm$ 3

Cultures of adult optic nerve cells were prepared as described previously for perinatal nerves², with two modifications; the incubations in trypsin and collagenase were increased to 30 min and the cells were washed in Dulbecco's modified Eagle's medium (DMEM) containing 10% FCS after the dissociation. Cells from one nerve (~50,000 live cells, as assessed by counting in a haemocytometer using Trypan blue) were plated in 20 μl of the same medium. In experiments using a poly-L-lysine (PLL) substrate, the cell suspension was placed on a PLL-coated coverslip in Falcon multi-well plates, left at 37 °C for 30–60 min to allow the cells to adhere and then 500 μl of DMEM supplemented as described previously³ (100 $\mu\text{g ml}^{-1}$ bovine serum albumin, 5 $\mu\text{g ml}^{-1}$ insulin, 100 $\mu\text{g ml}^{-1}$ transferrin, 62 ng ml^{-1} progesterone, 16.1 $\mu\text{g ml}^{-1}$ putrescine, 39.4 ng ml^{-1} selenium, 0.4 $\mu\text{g ml}^{-1}$ thyroxine, 0.3 $\mu\text{g ml}^{-1}$ triiodothyronine) as modified from Bottenstein and Sato²³ (B-S medium), containing 0.5% FCS (B-S/0.5% FCS), was added. In experiments using astrocyte conditioned medium, the cells were plated onto PLL-coated coverslips as described above, but after 30–60 min these coverslips were placed on glass spacers⁵ in a well containing 500 μl of B-S/0.5% FCS in which a monolayer of type-1 astrocytes, prepared as described previously⁵, had been growing and conditioning the medium for at least 18 h. In experiments using type-1 astrocyte monolayers as substrate, the 20- μl droplet was expelled from a pipette tip directly onto the astrocyte monolayer. All cultures were incubated for the time indicated at 37 °C in a humidified atmosphere of 5% CO_2 and 95% air. Cultures were labelled with A2B5 antibody as described in Fig. 1 legend and all antibody-labelled cells on the coverslip were counted. In autoradiography experiments, ^3H -thymidine (2.0 Ci mmol^{-1}) was added to the culture medium for the last 18 h, at a final concentration of 2 $\mu\text{Ci ml}^{-1}$. After labelling with A2B5 antibody, cultures were dried, dipped in Ilford K5 emulsion and exposed at -70 °C for 3 days before being developed in Ilford Contrast FF. Cells with ≥ 10 silver grains over the nucleus were counted as radiolabelled. These results and those in Table 2 are expressed as means $\pm$ s.d. of at least three separate experiments.

**Fig. 2** Autoradiograph of a pair of ^3H -thymidine-labelled, A2B5⁺ 0-2A progenitor-like cells from adult optic nerve growing on a monolayer of type-1 astrocytes. The cells were photographed using phase contrast (a) and rhodamine (b) optics. Note that both the 0-2A progenitor cells are radiolabelled; these cells were GFAP⁻ (not shown). Scale bar, 10 μm . The ^3H -thymidine (78 Ci mmol^{-1} , 2 $\mu\text{Ci g}^{-1}$) was injected into the heart of the rat 18 h and then 1 h before death. Optic nerve cells were cultured on type-1 astrocytes, double-labelled with A2B5 and anti-GFAP antibodies, processed for autoradiography as described for Table 1 and then exposed for 6 weeks before developing.

of eight times and differentiated into oligodendrocytes within 10 days or less⁶. In these single-cell experiments the proliferative capacity of individual 0-2A progenitor cells in newborn optic nerve varied widely, although the daughter cells of a single progenitor cell tended to divide a similar number of times before differentiating, even when separated and recultured individually on type-1 astrocyte monolayers in different microwells⁶. In contrast, in bulk co-cultures of perinatal optic nerve cells on type-1 astrocytes, new oligodendrocytes continue to be produced from dividing progenitor cells for weeks^{4,5}. Perhaps the simplest interpretation of these experiments is that 0-2A progenitor cells in perinatal optic nerve are produced continually from pre-0-2A progenitor cells. Taken together with the cell-division-counting model for the regulation of oligodendrocyte differentiation, this scheme would account for the heterogeneity in the proliferative capacity of 0-2A progenitor cells in nerves of a single age, as well as for the finding that oligodendrocytes are produced continually for at least several weeks after birth in the rat optic nerve^{15,16}. Our present findings suggest that the putative pre-0-2A progenitor cell produces proliferating 0-2A progenitor cells throughout life and, therefore, raise the possibility that this cell is a slowly dividing, self-renewing stem cell. If this is so, then

Table 2 Differentiation of putative 0-2A progenitor cells from adult optic nerve into either oligodendrocytes or type-2 astrocytes depending on the culture conditions

Culture medium	Total no. of cells per culture		
	Oligodendrocytes (GC ⁺)	Type-2 astrocytes (A2B5 ⁺ , GFAP ⁺)	Putative 0-2A progenitor cells (A2B5 ⁺ , GFAP ⁻ , GC ⁻)
Astrocyte conditioned medium (ACM) $\times$ 3 days	6 $\pm$ 7	44 $\pm$ 8	161 $\pm$ 47
ACM $\times$ 3 days then B-S/0.5% FCS $\times$ 3 days	150 $\pm$ 18	20 $\pm$ 13	19 $\pm$ 12
ACM $\times$ 3 days then B-S/10% FCS $\times$ 3 days	41 $\pm$ 30	206 $\pm$ 95	0 $\pm$ 12

Adult optic nerve cells were cultured in ACM for 3 days as described in Table 1 legend and then, in some cases, were switched to either B-S/0.5% FCS or to B-S/10% FCS for a further 3 days, after which they were labelled with A2B5 and anti-GFAP antibodies, or with anti-GC and A2B5 antibodies as described for Fig. 1. All antibody-labelled cells on the coverslip were counted and classified according to their antigenic phenotype. The number of putative 0-2A progenitor cells (A2B5⁺, GFAP⁻, GC⁻ process-bearing cells) was calculated after counting two separate coverslips from the same experiment, one stained with A2B5 and anti-GFAP and the other with anti-GC and A2B5 antibodies. The number of A2B5⁺, GC⁺ cells was subtracted from the number of A2B5⁺, GFAP⁻ cells, and the number of A2B5⁺, GFAP⁺ cells was subtracted from the number of A2B5⁺, GC⁻ cells; the mean of these two numbers was taken as the number of putative 0-2A progenitor cells in that experiment.

in this respect at least, 0-2A progenitor cells resemble rapidly dividing haematopoietic progenitor cells, which have a limited developmental and proliferative capacity and are thought to be derived from slowly dividing, self-renewing multipotential haematopoietic stem cells¹⁹.

The number of 0-2A progenitor cells in adult optic nerve is likely to be small. Although we do not know what proportion of these cells were liberated from the nerve by our dissociation procedures, or their plating efficiency, there were always less than 200 0-2A progenitor-like cells in cultures prepared from a single nerve and examined after 2 days. More importantly, in semi-thin frozen sections of adult rat optic nerve, less than 2% of the cells have the antigenic phenotype (A2B5⁺, GFAP⁻) of 0-2A progenitors².

The fate of the 0-2A progenitor cells in the adult optic nerve is uncertain. They may produce oligodendrocytes and/or type-2 astrocytes in association with a continuous slow turnover of these glial cells (as suggested for glial cells in other regions of the central nervous system (CNS)²⁰⁻²²), and/or in association with the continuous growth of the nerve during adult life. They may also produce new oligodendrocytes in response to demyelinating lesions. The relative inefficiency of remyelination in some demyelinating CNS diseases, such as multiple sclerosis, can have serious clinical consequences. Perhaps an understanding of the factors controlling the production, proliferation and differentiation of 0-2A progenitor cells in adult white matter will lead to the development of ways to enhance the production of new oligodendrocytes in these diseases.

We thank Dr E. Abney for performing the intracardiac injections of ³H-thymidine, Dr M. Nirenburg for the A2B5 hybridoma cells, Guus Wolswijk, Mark Noble, Sally Temple, Simon Hughes and other members of the Neuroimmunology group for helpful discussion. C.f.f.-C. is supported by a grant from the Multiple Sclerosis Society of Great Britain.

Received 1 October; accepted 3 December 1985.

1. Raff, M. C., Abney, E. A. & Miller, R. H. *Dev Biol.* **106**, 53-60 (1984).
2. Miller, R. H., David, S., Patel, R., Abney, E. A. & Raff, M. C. *Dev Biol.* **111**, 35-41 (1985).
3. Raff, M. C., Miller, R. H. & Noble, M. *Nature* **303**, 390-396 (1983).
4. Noble, M. & Murray, K. *EMBO J.* **3**, 2243-47 (1984).
5. Raff, M. C., Abney, E. A. & Fok-Seang, J. *Cell* **42**, 61-69 (1985).
6. Temple, S. & Raff, M. C. *Cell* (in the press).
7. Eisenbarth, G. S., Walsh, F. S. & Nirenburg, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4913-17 (1979).
8. Bignami, A., Eng, L. F., Dahl, D. & Uyeda, C. T. *Brain Res.* **43**, 429-435 (1972).
9. Raff, M. C. *et al. Brain Res.* **174**, 283-308 (1979).
10. Ranscht, B., Clapham, P. A., Price, J., Noble, M. & Seifert, W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2709-2713 (1982).
11. Raff, M. C. *et al. Nature* **274**, 813-816 (1978).
12. David, S., Miller, R. H., Patel, R. & Raff, M. C. *J. Neurocytol.* **13**, 961-974 (1984).
13. Temple, S. & Raff, M. C. *Nature* **313**, 223-225 (1985).
14. Raff, M. C., Williams, B. P. & Miller, R. H. *EMBO J.* **3**, 1857-1864 (1984).
15. Skoff, R. P., Price, D. L. & Stocks, A. J. *comp. Neurol.* **169**, 291-312 (1976).
16. Skoff, R. P., Price, D. L. & Stocks, A. J. *comp. Neurol.* **169**, 313-334 (1976).
17. Miller, R. H. *et al. J. Neurosci.* (in the press).
18. Abney, E. A., Williams, B. P. & Raff, M. C. *Dev Biol.* **100**, 166-171 (1983).
19. Till, J. E. & McCulloch, E. A. *Biochim. biophys. Acta* **605**, 431-459 (1980).
20. Mares, V., Lodin, Z. & Jilek, M. J. *comp. Neurol.* **161**, 471-482 (1975).
21. Kaplan, M. S. & Hinds, J. W. *J. comp. Neurol.* **193**, 711-727 (1980).
22. Lindsay, R. M. *Adv. Cell Neurobiol.* (in the press).
23. Bottenstein, J. & Sato, G. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 514-517 (1979).
24. Pruss, R. M. *Nature* **280**, 688-690 (1979).
25. Johnson, G. D. *et al. J. immun. Meth.* **55**, 231-242 (1982).

Restricted recognition of β_2 -microglobulin by cytotoxic T lymphocytes

Hans-Georg Rammensee*†, Peter J. Robinson‡, Andrea Crisanti* & Michael J. Bevan†

* Basel Institute for Immunology, Grenzacherstrasse 487, Postfach, 4005 Basel, Switzerland

† Department of Immunology, Scripps Clinic & Research Foundation, La Jolla, California 92037, USA

‡ Institute for Immunology and Genetics, German Cancer Research Center, Im Neuenheimer Feld 280, D-6900, Heidelberg, FRG

Recognition of foreign antigen by cytotoxic T lymphocytes (CTL) is restricted by class I major histocompatibility complex (MHC) products. Class I heavy chains (relative molecular mass (M_r) 45,000-48,000) are reversibly and noncovalently associated with β_2 -microglobulin (β_2M , $M_r = 12,000$) (see ref. 1). Cells expressing human or murine class I heavy chains can exchange their native β_2M for exogenously added free β_2M (refs 2-4), which is present in serum⁵⁻⁷. Two allelic forms of β_2M exist among the common laboratory mouse strains, β_2M -A and β_2M -B, which are represented in BALB and C57BL mice, respectively⁸⁻¹⁰. The two forms differ at a single amino acid at position 85 (ref. 11), the gene (β_2m) is located on chromosome 2 linked to a minor histocompatibility (H) region, $H-3$ (refs 12, 13). It has been proposed that one of the $H-3$ loci is identical with β_2m , and that CTL raised across certain $H-3$ incompatibilities are actually specific for β_2M ^{14,15}. Here we describe CTL raised in such a combination which recognize endogenous as well as exogenous β_2M -B in the context of $H-2K^b$. This represents a unique case of CTL recognition, as CTL usually recognize antigens inserted into the membrane, and it is the first molecular identification of the product of a minor H locus.

The congenic resistant strains C57BL/10 and the partner strain B10.C-H-3^c differ at the minor H region $H-3$ and the

Table 1 β_2M -B-specific antibodies inhibit killing by 181B

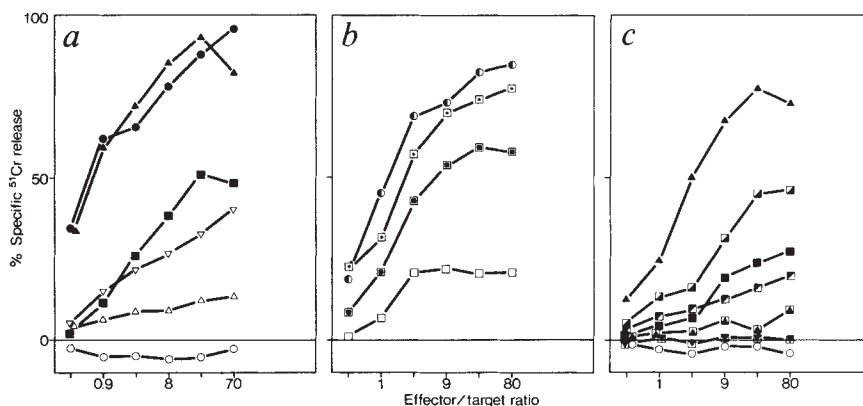
Antibody	Dilution	Specific ⁵¹ Cr release of C57BL/6 blasts (%)
None	—	74.8
28-11-5S (anti-D ^b) (ref. 34)	1/60	70.0
	1/300	83.0
5F1.2 (anti-K ^b) (ref. 35)	1/60	40.3
	1/300	55.7
	1/1,500	67.8
S19.8 (anti- β_2M -B)	1/60	9.7
	1/300	15.2
	1/1,500	19.8

⁵¹Cr-labelled C57BL/6 blasts were incubated for 4 h with the indicated antibodies^{34,35} (as ascites fluid) and 181B CTL at an effector/target ratio of 38/1. Other conditions are as described in Fig. 1.

β_2m locus on chromosome 2 (refs 12, 13, 16 and H.G.R., unpublished data). In B10.C-H-3^c these genes are derived from the BALB genome. The two strains reject each other's skin grafts¹⁶ and produce H-2K^b-restricted CTL against each other on *in vivo* immunization¹⁵. It has been proposed that the $H-3$ region consists of several H loci, one of them being β_2m ^{14,15,17,18}. To test the hypothesis of identity between one of the $H-3$ products and β_2M , we produced a CTL line from (BALB.K × B10.C-H-3^c)F₁ ($H-2K^b$, β_2m^a) mice immunized with (BALB.K × C57BL/10)F₁ ($H-2K^b$, $\beta_2m^{a \times b}$). (We used $H-2K^b$ heterozygous animals in the hope of obtaining H-2K^b-restricted β_2M -B-specific CTL, which could have been tested on L929 cells transfected with the β_2m^b gene¹⁹.)

The CTL line (181B) killed stimulator-type targets and also C57BL/10 ($H-2^b$, β_2m^b), C57BL/6 ($H-2^b$, β_2m^b) and B10.A(5R) ($H-2K^bD^k$, β_2m^b), but not responder-type targets nor BALB.B ($H-2^b$, β_2m^a), B10.BR ($H-2^k$, β_2m^b) or B10.A (2R) ($H-2K^dD^b$, β_2m^b) (Fig. 1 and H.G.R., unpublished data), suggesting that it is specific for β_2M -B and is H-2K^b-restricted.

Fig. 1 Characterization of the antigen recognized by the CTL line 181B [(BALB.K×B10.C-H-3^c)F₁ (H-2^{k×b}, β_2m^a) anti-(BALB.K×C57BL/10)F₁ (H-2^{k×b}, $\beta_2m^{a×b}$)]. CTL were tested on ⁵¹Cr-labelled blasts of responder (indicated as CK×H-3^c) or stimulator (CK×B10) type or of C57BL/6 (B6; β_2m^b) or BALB.B (CB; β_2m^b) origin. Blasts were induced for 2 days with mitogen and C57BL serum (B1 serum) or BALB serum (C serum) at 5% concentration (if not indicated otherwise) or with 10% FCS. **a**, H-2^b, β_2m^a cells can acquire the antigen recognized by 181B CTL from C57BL, but not BALB serum. BALB.B blasts were generated in 5% C57BL serum which had been absorbed with β_2m -B-specific monoclonal antibodies S19.8 bound to protein A-Sepharose or on protein A-Sepharose without antibody. The assay was performed in Iscove's medium containing 2% bovine serum albumin (BSA) and 1% BALB serum. Spontaneous release of targets was between 24.7 and 29.1%. ●, B6, B1 serum; ▲, B6, C serum; ■, CB, B1 serum; ▽, CB, B1 serum, protein A absorbed; △, CB, B1 serum and S19.8; ○, CB, C serum. **b**, Maximal killing of H-2^b, β_2m^a blasts with acquired serum-born antigen can be demonstrated under optimal conditions. This assay was performed in RPMI 1640 medium containing 2% bovine serum albumin (BSA) and 2.5% C57BL serum. Spontaneous release ranged between 24.2 and 37.8%. ●, CK×B10, FCS; □, CK×H-3^c, 5% B1 serum; ▤, CK×H-3^c, 1.25% B1 serum; □, CK×H-3^c, FCS. **c**, The CTL target antigen acquired by H-2^b, β_2m^a cells is β_2m -B. BALB.B blasts were raised in serum-free Iscove's medium containing 2% BSA and 200, 40 or 8 ng ml⁻¹ purified β_2m -B or 200 ng ml⁻¹ purified β_2m -A. Assay in 1% BALB serum; spontaneous release between 19.4 and 34.3%. ▲, B6, C serum; ■, CB, 200 ng β_2m -B; ▤, CB, 40 ng β_2m -B; ▥, CB, 8 ng β_2m -B; ▦, CB, 200 ng β_2m -A; ○, CB, C serum.



The ability of anti- β_2m -B and anti-H-2K^b, but not anti-H-2D^b, antibodies to inhibit lysis by 181B is consistent with this view (Table 1).

That the line is specific for β_2m -B is shown by the following experiments. BALB.B lymphoblasts raised in medium containing 5% C57BL serum (which is known to contain free β_2m -B²⁰) instead of the usual fetal calf serum (FCS) were killed by 181B, but BALB.B blasts raised in 5% BALB serum were not killed (Fig. 1a). C57BL/6 or B10.BR sera were both effective, sera of BALB/c, BALB.B or BALB.K were all ineffective (data not shown). The presence of β_2m -B on the surface of β_2m^a blasts raised with β_2m -B-containing serum was demonstrated by the binding of radiolabelled β_2m -B-specific monoclonal antibodies S19.8 (ref. 21; Table 2). Thus, sera containing β_2m -B, but not β_2m -A, can make BALB.B targets susceptible to lysis by 181B CTL, irrespective of the H-2 haplotype of the serum donor. BALB.B blasts raised in C57BL serum absorbed with β_2m -B-specific S19.8 antibodies bound to protein A-Sepharose were only weakly killed. Absorption of C57BL serum on protein A-Sepharose alone did not significantly reduce its capacity to supply the target structure to the BALB.B cells (Fig. 1a). Blasts raised in 1.25% C57BL serum were killed less efficiently than blasts raised in 5% serum (Fig. 1b). When β_2m^a targets raised in C57BL serum were tested in the killer assay also containing C57BL serum, the level of lysis was similar to that of β_2m^b targets (Fig. 1b). β_2m^a targets raised in BALB serum were also lysed at a low level if the assay was run in C57BL serum, probably through β_2m^a targets picking up free β_2m -B during the 4-h assay⁶.

To identify unequivocally our serum-born CTL antigen, we tried to substitute the native β_2m -A on BALB.B cells with purified urinary β_2m (ref. 22). BALB.B blasts raised in serum-free medium containing 200 ng ml⁻¹ purified β_2m -B were killed efficiently by 181B CTL, and lower concentrations of β_2m -B yielded lower killing (Fig. 1c). BALB.B blasts raised in 200 ng ml⁻¹ purified β_2m -A were not lysed at all. Thus, killing compar-

Table 2 Serum β_2m -B binds to the surface of β_2m^a cells

Cells	Serum source	Cell no. (×10 ⁶)	³ H-labelled S19.8 antibodies bound (c.p.m.)
C57BL/6 (β_2m^b)	FCS	1.5	16 518
		0.5	8 156
B10.C-H-3 ^c (β_2m^a)	FCS	1.5	114
		0.5	93
	B10.BR	1.5	3 875
		0.5	2 512

Spleen cells of the indicated strains were cultured for 2 days in medium containing either 10% FCS or 5% B10.BR serum. The binding assay was performed as described elsewhere³⁶. Briefly, cells were pelleted, incubated for 60 min with ³H-labelled β_2m -B-specific S19.8 antibodies and washed twice.

able to that obtained by acquired serum-born antigen was demonstrated with purified β_2m -B, but not β_2m -A.

In conclusion, the CTL line 181B recognizes β_2m -B in the context of K^b. This represents a unique case of CTL recognition, because 181B recognizes a soluble antigen in association with class I heavy chain. CTL usually recognize membrane-inserted antigens, that is, antigenic epitopes on class I heavy chains or class I plus virus-encoded protein; or class I plus hapten covalently coupled to cell-surface antigen; or class I plus allogeneic cell-surface antigens²³.

Of interest is the density of β_2m on the cell surface. It is well known that the reactivity of lymphocytes to a given antigen is dependent on the density of that antigen on the presenting cell²⁴⁻²⁶. On this basis it was proposed that the strong alloreactivity to foreign MHC products is caused by the high density

of MHC on antigen-presenting cells, as opposed to the density of complexes of antigen plus MHC²⁷. Because the density of β_2 M on stimulator cells is at least as high as that of K and D heavy chains together, one would expect a high frequency of β_2 M-reactive cells in unprimed animals, comparable to that of alloreactive cells. This is found not to be the case: no primary CTL response between MHC-identical, β_2 M-incompatible strains can be demonstrated²⁸ and foreign β_2 M does not contribute to the CTL response between class I-incompatible strains²⁹. We conclude that antigen density alone cannot account for the high frequency of alloreactive T cells, and that the region of the MHC antigen affected by a change is also essential.

Our findings support the earlier hypothesis that β_2 M is one of the minor H antigens encoded in the H-3 region. Characteristics of minor H antigens are: (1) they cause slow rejection of grafts³⁰; (2) they elicit CTL responses only after *in vivo* priming²⁸; and (3) their recognition by CTL is class I restricted^{31,32}. We have shown that the latter two points apply to β_2 M; whether β_2 M incompatibility can lead to graft rejection remains to be demonstrated. No other minor H antigen has so far been identified on the molecular level.

We thank U. Hämmerling, K. Fischer Lindahl and L. Sherman for antibodies, Silvie Bove for technical assistance, K. Fischer Lindahl and W. Haas for improving, and Judie Hossmann and Carolyn Harley for typing the manuscript. H.G.R. was supported by the Deutsche Forschungsgemeinschaft during part of the work. The Basel Institute for Immunology was founded and is supported by F. Hoffman-La Roche and Co., Basel, Switzerland. *Note added in proof:* Since this manuscript was submitted, Kurtz *et al.*³⁷ have also reported that CTL raised across certain H-3/ β_2 m incompatibilities can be blocked by β_2 M-specific antibodies.

Received 30 July; accepted 12 December 1985.

- Klein, J. *Science* **203**, 516-521 (1979).
- Hyafil, F. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5834-5838 (1979).
- Schmidt, W., Festenstein, H., Ward, P. J. & Sanderson, A. R. *Immunogenetics* **13**, 483-491 (1981).
- Hester, R. B., Kubo, R. T. & Grey, H. M. *Scand. J. Immun.* **9**, 125-134 (1979).
- Kefford, R. F., Calabi, F., Fearnly, I. M., Burrone, O. R. & Milstein, C. *Nature* **308**, 641-642 (1984).
- Bernabeu, C., van de Rijn, M., Lerch, P. G. & Terhorst, C. P. *Nature* **308**, 642-645 (1984).
- Berggard, I. & Bearn, A. G. *J. biol. Chem.* **243**, 4095-4103 (1968).
- Goding, J. & Walker, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7395-7399 (1980).
- Michaelson, J., Rothenberg, E. & Boyse, E. A. *Immunogenetics* **11**, 93-95 (1980).
- Robinson, P. J., Graf, L. & Sege, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1167-1170 (1981).
- Gates, F. T. III, Coligan, J. E. & Kindt, T. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 554-558 (1981).
- Michaelson, J. *Immunogenetics* **13**, 167-171 (1981).
- Robinson, P. J. *et al. Immunogenetics* **14**, 449-452 (1981).
- Michaelson, J. *Immunogenetics* **17**, 219-259 (1983).
- Rammensee, H.-G. & Klein, J. *J. Immun.* **130**, 2926-2929 (1983).
- Graf, R. J., Brown, D. & Snell, G. D. *Transplant. Proc.* **5**, 299-302 (1973).
- Roopenian, D. C. & Click, R. E. *Immunogenetics* **10**, 333-341 (1980).
- Click, R. E., Schneider, D. & Roopenian, D. C. *J. Immun.* **126**, 2378-2381 (1981).
- Margulies, D. H., Parnes, J. R., Johnson, N. A. & Seidman, J. G. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2328-2331 (1983).
- Kimura, S., Tada, N., Lin, Y. & Hämmerling, U. *Immunogenetics* **14**, 445-447 (1981).
- Tada, N., Kimura, S., Hatzfeld, A. & Hämmerling, U. *Immunogenetics* **11**, 441-447 (1980).
- Graf, L., Peterson, P. A., Sege, K. & Robinson, P. J. *Molec. Immun.* **19**, 877-884 (1982).
- Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51-177 (1979).
- Pohlitz, H., Haas, W. & v. Boehmer, H. *Eur. J. Immun.* **9**, 681-690 (1979).
- Matis, L. M., Glimcher, L. H., Paul, W. E. & Schwartz, R. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6019-6023 (1983).
- Janeway, C. A. *et al. Nature* **306**, 80-82 (1983).
- Bevan, M. J. *Immun. Today* **5**, 128-130 (1984).
- Bevan, M. J. *J. exp. Med.* **143**, 1283-1288 (1979).
- Langhorne, J. & Fischer Lindahl, K. *Eur. J. Immun.* **12**, 101-106 (1982).
- Graff, R. J. in *Origins of Inbred Mice* (ed. Morse, H. C. III) 371-388 (Academic, New York, 1978).
- Bevan, M. J. *J. exp. Med.* **142**, 1349-1364 (1975).
- Gordon, R. D., Simpson, E. & Samelson, L. E. *J. exp. Med.* **142**, 1108-1120 (1975).
- Rammensee, H.-G., Fink, P. J. & Bevan, M. J. *J. Immun.* **133**, 2390-2396 (1984).
- Ozato, K. & Sachs, D. H. *J. Immun.* **126**, 317-321 (1981).
- Rüsch, E., Kuon, W. & Hämmerling, G. J. *Transplant. Proc.* **15**, 2093-2096 (1983).
- Fischer Lindahl, K. *Immun. Meth.* **3**, 187 (1985).
- Kurtz *et al. J. Immun.* **135**, 2847 (1985).

Two roles for guanine nucleotides in the stimulus-secretion sequence of neutrophils

M. M. Barrowman, S. Cockcroft & B. D. Gomperts

Department of Experimental Pathology, University College London, University Street, London WC1E 6JJ, UK

The term 'stimulus-secretion coupling' has, since first enunciated¹, been held to involve the mobilization of cytosol Ca^{2+} , which in turn is sufficient to trigger exocytotic secretory processes in metabolically competent cells. However, recent studies on a wide range of secretory cell types indicate that a role for Ca^{2+} can be obviated: examples are stimulation with phorbol ester (phorbol myristate acetate, PMA)^{2,3} which causes the activation of protein kinase C⁴ or the stimulation of platelets with collagen⁵. Ca^{2+} -independent exocytosis also occurs when analogues of GTP are injected through the lumen of patch pipettes directly into the cytosol of mast cells⁶. The results presented here suggest that GTP analogues can activate secretory processes by actions at two distinct locations: one may be at the level of the receptor⁷ involving the activation of polyphosphoinositide (PPI) phosphodiesterase⁸ with consequent liberation of diacylglycerol (DG)⁹; the other involves direct activation of the exocytotic mechanism. These conclusions are based on measurements of exocytotic secretion from permeabilized neutrophils into which we have been able to introduce, individually and in combination, Ca^{2+} chelators (EGTA and BAPTA), Ca^{2+} (buffered at micromolar concentrations with EGTA), analogues of GTP and GDP and the direct activator of protein kinase C, PMA.

To permeabilize rabbit neutrophils, we used 'three-day, egg-grown' Sendai virus¹⁰. The virus, which binds at the ubiquitous sialic acid groupings of either glycoproteins¹¹ or glycolipids¹² present on the surfaces of all animal cells, causes leakage of intracellular metabolites¹³ and conversely it is possible to intro-

duce normally impermeant solutes into the cytosol of such cells¹⁴. We have previously used Sendai virus to permeabilize rat peritoneal mast cells and introduce Ca^{2+} buffers to control cytosol Ca^{2+} at concentrations in the micromolar range: this caused the cells to undergo exocytotic secretion of histamine and N-acetylglucosaminidase¹⁵.

Introduction of Ca^{2+} (buffered at concentrations in the micromolar range) into permeabilized neutrophils causes exocytotic secretion of β -glucuronidase (Fig. 1a, lowest curve): half-maximal secretion is achieved at $p\text{Ca}$ 6.3 [0.56 ± 0.1 M ($\pm$ s.d., $n = 17$ experiments)] and is independent of the concentration of EGTA (in the range 0.3-10 mM) used to buffer the Ca^{2+} . The mean extent of maximal secretion was $31 \pm 14\%$ ($\pm$ s.d., $n = 57$). As with the secretion from intact cells due to receptor-directed ligands, secretion from the permeabilized neutrophils also requires brief pretreatment with cytochalasin B. There is in addition, as with other permeabilized cell systems including bovine adrenal cells¹⁶, platelets¹⁷ and pancreatic islet β -cells¹⁸, a requirement for metabolic support which must be provided in the form of Mg-ATP¹⁹. In distinction to some of the other systems cited, we have found that the sensitivity of β -glucuronidase secretion to Ca^{2+} and other stimulatory agents (see below) declines for some 10-15 min following permeabilization with Sendai virus. This is not due to the closure of the lesions¹⁹.

In view of the observation²⁰ that analogues of GTP can spare the requirement of Ca^{2+} for serotonin secretion from permeabilized platelets and so permit secretion to occur at concentrations of Ca^{2+} down to 10^{-8} M, we tested the effects of GTP- γ -S on Ca^{2+} -induced β -glucuronidase secretion from permeabilized neutrophils (Fig. 1a). With these cells, the effects of Ca^{2+} and the GTP analogue are additive and independent of each other. The permeabilized cells loaded with GTP- γ -S show no requirement for Ca^{2+} in the secretory process. This conclusion must of course be tempered by one's understanding of the term 'no Ca^{2+} ' (ref. 21). If we accept the inescapable presence of approxi-

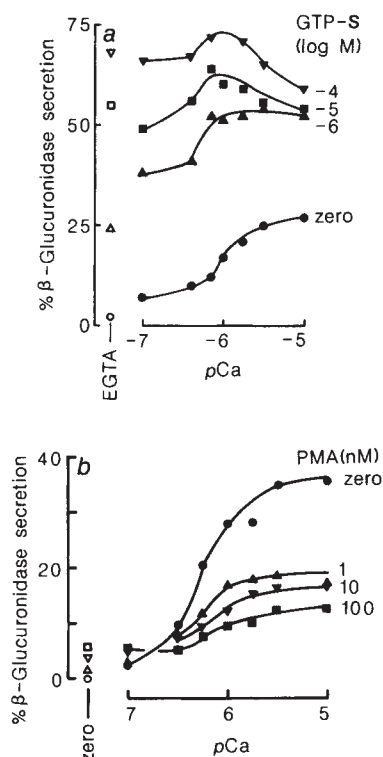


Fig. 1 Dependence on Ca^{2+} of β -glucuronidase secretion from permeabilized rabbit neutrophils in the presence of GTP- γ -S (a) and phorbol myristate acetate (b). a, Rabbit neutrophils were permeabilized with Sendai virus and then treated with Ca^{2+} at the concentrations indicated in the absence (●) and presence of GTP- γ -S at 10^{-6} M (▲), 10^{-5} M (■) and 10^{-4} M (▼). b, Rabbit neutrophils were permeabilized as in a and then transferred to solutions containing Ca^{2+} buffered at the concentrations indicated in the absence (●) or presence of PMA at 1 (▲), 10 (▼) and 100 (■) nM.

Methods. Rabbit peritoneal exudate neutrophils were isolated as previously described³⁸ and suspended at $2.5 \times 10^5 \text{ ml}^{-1}$ in a buffer (pH 6.8, comprising 137 mM NaCl, 2.7 mM KCl, 2 mM MgCl_2 , 20 mM HEPES, 5.6 mM glucose, 15 μM EGTA and 1 mg ml^{-1} bovine serum albumin). After incubation for 20 min at 37°C , cytochalasin B (5 $\mu\text{g ml}^{-1}$ final concentration) was added together with Mg-ATP (final 5 mM) and EGTA 0.6 mM. After 2.5 min, Sendai virus (100 haemagglutinating units (HU) per ml final concentration) was added and after a further 2.5 min the cells were diluted twofold in tubes containing Ca^{2+} (buffered with 3 mM EGTA: final concentration) at the concentrations indicated and GTP- γ -S. b, The formulations for the pCa^{2+} buffers were calculated by use of the program CHELATE using appropriate dissociation constants for 37°C derived from values of ΔH and ΔS (ref. 39). After 10 min incubation, cells were sedimented by centrifugation at 4°C and samples of the supernatant were withdrawn for analysis of β -glucuronidase²² and lactate dehydrogenase¹⁵. Leakage of lactate dehydrogenase was measured on many occasions during this work and in no instance did it exceed 5%. All data points indicate averages of duplicate determinations. Similar results were obtained on at least three occasions.

mately $10 \mu\text{M}$ of contaminating Ca^{2+} (mainly present in the bovine serum albumin, a component of the buffered salt solution²²), then under the routine conditions of our present experiments (pH 6.8 and 3 mM EGTA) we calculate an equilibrium concentration of 'free' Ca^{2+} of 2×10^{-9} M. We have also carried out similar experiments under conditions of elevated pH (7.4) and EGTA (30 mM) or BAPTA (30 mM) (so reducing Ca^{2+} to $< 2 \times 10^{-11}$ M). In this case we find (Fig. 2) that the extent of secretion due to GTP- γ -S is increased, not decreased, as the concentration of either chelator is increased from 3 to 30 mM while leakage of lactate dehydrogenase remains below 5%. We cannot explain why EGTA supports a higher level of secretion

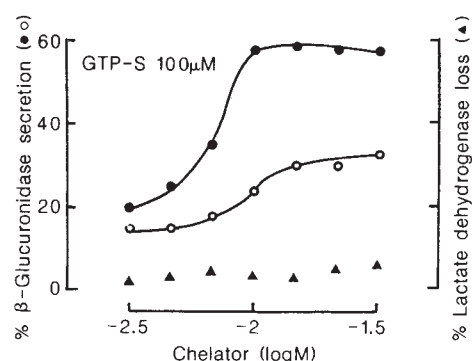


Fig. 2 Effect of extensive Ca^{2+} depletion on GTP- γ -S-induced secretion of β -glucuronidase from permeabilized neutrophils. Rabbit neutrophils (prepared as for Fig. 1) were treated with EGTA (●) or BAPTA (○) at the indicated concentrations for 2.5 min and then permeabilized by addition of Sendai virus (100 HU ml^{-1}). GTP- γ -S ($100 \mu\text{M}$) was added after a further 2.5 min and the cells were further incubated and then processed as described in Fig. 1 legend for measurement of β -glucuronidase secretion. Solid triangles indicate release of lactate dehydrogenase in the EGTA experiment. Similar results were obtained on three occasions.

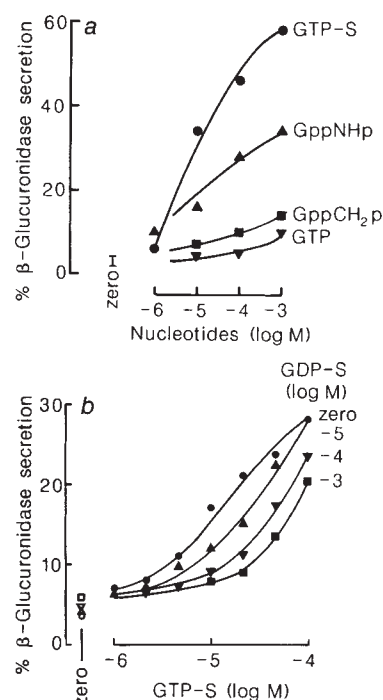


Fig. 3 Stimulation of Ca^{2+} -independent secretion by analogues of GTP and its inhibition by GDP- β -S. Rabbit neutrophils (prepared as described in Fig. 1 legend) were treated with 3 mM EGTA for 2.5 min. They were then permeabilized with Sendai virus and after a further 2.5 min, they were processed as follows: a, the cells were transferred to solutions containing 3 mM EGTA and GTP analogues at the concentrations indicated: GTP- γ -S (●), GppNHp (▲), GppCH₂p (■), GTP (▼). Similar results were obtained on three occasions. b, The cells were transferred to tubes containing GTP- γ -S and GDP- β -S at the concentrations indicated: zero GDP- β -S (●), 10^{-5} M (▲), 10^{-4} M (▼), 10^{-3} M (■). Similar results were obtained on four occasions. Further processing was as described in Fig. 1 legend.

than BAPTA at all concentrations. It should further be stressed that the guanine nucleotide has always been applied to the cells 2.5 min after permeabilization and introduction of the chelate ligand. We would claim that biological processes occurring under such conditions can justly be regarded as Ca^{2+} -independent.

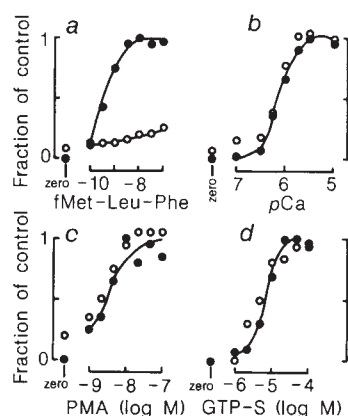


Fig. 4 Effect of pertussis toxin on secretion from intact neutrophils stimulated by fMet-Leu-Phe, and from permeabilized neutrophils stimulated by phorbol ester, Ca^{2+} or GTP- γ -S. Rabbit neutrophils were incubated in a balanced salt solution (containing 1 mM Ca^{2+}) with pertussis toxin (500 ng ml $^{-1}$) for 90 min. They were then processed as follows: *a*, cells were washed twice and suspended in a solution containing 1 mM Ca^{2+} and 5 $\mu\text{g ml}^{-1}$ cytochalasin B and then stimulated by addition of fMet-Leu-Phe at the concentrations indicated. *b*, Cells were washed twice in a solution containing 0.6 mM EGTA and then treated with Sendai virus (100 HU ml $^{-1}$); after 2.5 min they were transferred to tubes containing Ca^{2+} buffered at the concentrations indicated. *c*, *d*, Cells were washed twice in a solution containing 3 mM EGTA and then further incubated for 2.5 min in 3 mM EGTA before addition of Sendai virus; they were then transferred to tubes containing phorbol ester at the indicated concentrations (*c*) or GTP- γ -S (*d*). All cells were then processed as described in Fig. 1 legend for measurement of β -glucuronidase secretion. Solid symbols (●) indicate control cells, open symbols (○) indicate pertussis-treated cells. The extent of maximal secretion due to fMet-Leu-Phe was $64 \pm 11\%$ ($\pm$ s.d., $n = 14$) and 61% in the experiment illustrated (*a*); Ca^{2+} , $31 \pm 14\%$ ($\pm$ s.d., $n = 57$) and 45% in the experiment illustrated (*b*); PMA (100 nM), $15 \pm 10\%$ ($\pm$ s.d., $n = 18$) and 20% in the experiment illustrated (*c*); GTP- γ -S (100 μM) $35 \pm 13\%$ ($\pm$ s.d., $n = 36$) and 32% in the experiment illustrated (*d*). All incubations were for 10 min. Similar results were obtained on four occasions.

We tested other commercially available analogues of GTP for the ability to stimulate Ca^{2+} -independent secretion of β -glucuronidase (Fig. 3*a*). Of these, GppNHp induces $\sim 60\%$ of that achieved by GTP- γ -S at all concentrations over the range 10^{-5} – 10^{-3} M. There was a small amount of secretion with GppCH $_2$ p ($< 25\%$) and GTP. Purified GTP (provided by F. Eckstein) is somewhat more active than the commercial material. GDP and GDP- β -S do not support secretion but are inhibitory of the secretion due to the GTP analogues. As the concentration of the inhibitory GDP (analogue) is increased, a higher concentration of the activator (GTP- γ -S) is needed to achieve the same amount of secretion in the manner characteristic of competitive inhibition (Fig. 3*b*).

In the knowledge⁸ that GTP- γ -S can activate the PPI phosphodiesterase of isolated neutrophil membranes in the presence of normal resting levels of Ca^{2+} , we tested the ability of GTP- γ -S to cause PPI breakdown in the presence of 3 mM EGTA and no added Ca^{2+} (that is, the conditions pertaining in the permeabilized cell experiments described in this paper). In four experiments GTP- γ -S (100 μM) selectively activated the hydrolysis of $19 \pm 1\%$ of phosphatidylinositol biphosphate and thus it is at least plausible that GTP- γ -S-induced secretion is due to its ability to activate the PPI phosphodiesterase. While it is unlikely that the water-soluble product inositol 1,4,5-trisphosphate could cause a rise of Ca^{2+} in the cytosol of the permeabilized cells to concentrations commensurate with the secretory process, the lipid-soluble product DG could be active in stimulating protein kinase C.

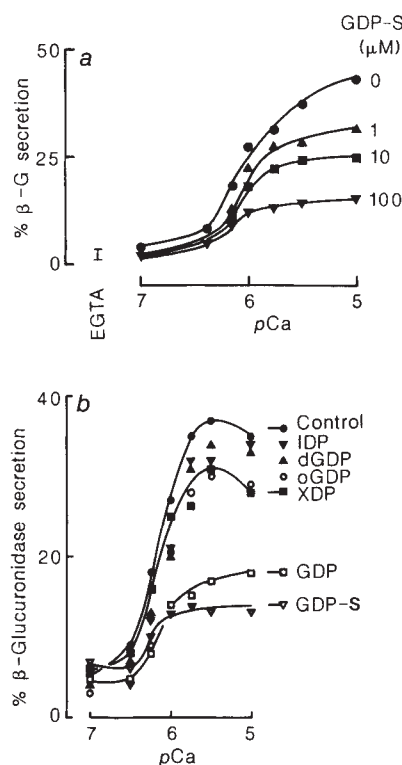


Fig. 5 Effect of GDP and GDP analogues on secretion due to Ca^{2+} from permeabilized rabbit neutrophils. Cells were prepared and permeabilized as described in Fig. 1 legend. *a*, Cells were transferred to solutions containing Ca^{2+} buffered at the concentrations indicated either in the absence (●) or in the presence of GDP- β -S at 10^{-6} M (▲), 10^{-5} M (■) and 10^{-4} M (▼). Similar results were obtained on four occasions. *b*, Cells were transferred to solutions containing Ca^{2+} buffered at the concentrations indicated in the absence (●) or in the presence of IDP (▼), dGDP (▲), oGDP (○), XDP (■), GDP (□) and GDP- β -S (▽) all at 10^{-4} M. Similar results were obtained on three occasions.

Protein kinase C activation can occur at concentrations of Ca^{2+} at least as low as 10^{-8} M (refs 3, 23) and we have found that secretion can occur on treating neutrophils with PMA (Fig. 4*c*) under conditions of vanishingly low concentrations of Ca^{2+} . PMA inhibited secretion due to Ca^{2+} (Fig. 1*b*): PMA (100 nM) caused $41 \pm 13\%$ ($\pm$ s.d., $n = 21$) inhibition of the secretion due to 10^{-5} M Ca^{2+} . This contrasts with the combination Ca^{2+} and GTP- γ -S (Fig. 1*a*) which causes additive effects on secretion under conditions in which DG must be generated. This indicates the possibility that GTP- γ -S may act at a site insensitive to any inhibitory effect of PMA on Ca^{2+} -induced exocytosis. An alternative explanation for the difference between PMA and GTP- γ -S on Ca^{2+} -induced secretion may arise from the transient nature of the DG signal due to the guanine nucleotide compared with the persistent activation of protein kinase C due to PMA. Yet another possibility is that the inhibitory action of PMA is not expressed through activation of protein kinase C.

G-proteins can generally be modified by ADP ribosylation induced by toxins from cholera^{24,25} and pertussis²⁶. The neutrophil responses to fMet-Leu-Phe of Ca^{2+} mobilization, PPI hydrolysis in both intact cells²⁷ and plasma membranes²⁸, and later responses, including secretion^{29,30}, chemotaxis²⁹ and the respiratory burst²⁹, are all suppressed in pertussis toxin-treated cells³¹. In order to distinguish the effect of GTP- γ -S in exocytosis from its role in receptor-mediated events, we examined the effect of pertussis treatment on the secretion due to fMet-Leu-Phe from intact cells and Ca^{2+} buffers, PMA and GTP- γ -S from permeabilized cells. The results indicate that whilst secretion

due to the receptor-directed ligand is prevented in toxin-treated cells, there is no effect on the secretion due to Ca^{2+} , PMA or GTP- γ -S (Fig. 4). We conclude that there are two sites of action of GTP in the stimulus-secretion sequence.

Further evidence in favour of the idea of a second site for GTP in the activation of exocytosis comes from the finding of inhibition by GDP and GDP- β -S on Ca^{2+} -induced secretion. Figure 5a illustrates the results of an experiment in which GDP- β -S was tested as a modulator of secretion of β -glucuronidase due to Ca^{2+} . GDP is also inhibitory to Ca^{2+} -induced secretion but higher concentrations are required (maximal inhibition at 1 mM; results not shown). Inhibition due to GDP is specific: the structurally related nucleoside diphosphates dGDP and IDP are without effect on Ca^{2+} -induced secretion (Fig. 5b). XDP and periodate-oxidized GDP caused inhibition to a lesser extent than GDP.

The results presented here strongly suggest that GTP and its analogues act at two sites in the control of the stimulus-secretion process. One of these has already been identified: this is the PPI phosphodiesterase⁸. The location of the second site is not known but it is concerned with the exocytotic mechanism as argued below.

GTP- γ -S can induce secretion and activate PPI phosphodiesterase in the absence of Ca^{2+} . If the sole effect of GTP- γ -S is to generate DG and so activate protein kinase C, then PMA should mimic GTP- γ -S. However, PMA only mimics GTP- γ -S in respect of inducing Ca^{2+} -independent exocytosis. The two differ in that secretion due to Ca^{2+} is inhibited by PMA whereas it is amplified by GTP- γ -S; furthermore, the extent of secretion due to PMA is always less than that due to GTP- γ -S (see Fig. 4 legend). Therefore, there is a second site of action at which GTP analogues cause secretion without intervention of the products of PPI-phosphodiesterase activation. There is the further possibility that Ca^{2+} -dependent secretion is also subject to regulation by guanine nucleotides. (Ca^{2+} -dependent secretion cannot be attributed to activation of PPI phosphodiesterase with consequent generation of DG, because the concentrations of Ca^{2+} involved are far too low to activate the enzyme⁹.) The finding of inhibition of Ca^{2+} -dependent secretion by GDP and GDP- β -S further supports the idea of a GTP-regulatory system at the exocytotic site.

Although secretion from neutrophils can be activated in a variety of ways (some of which are exemplified in the present paper), the mechanism of exocytosis is likely always to be the same. Central to the mechanism of exocytosis is the fusion of the membranes of the secretory granules with the plasma membrane³²⁻³⁴, the precise details of which are shrouded in obscurity. It is clear, however, that the main energy barrier to the close approach (within 2 nm) and then fusion of membranes lies in the forces which act to maintain the hydration of the membrane surfaces³⁵. These are enormous and it is therefore likely that the membrane fusions which occur in the exocytotic process are catalytically regulated. While none of the molecular events are known, the second site for GTP regulation must surely be at this general level because, regardless of the mode of activation of neutrophils, GDP and GDP- β -S are inhibitory (Ca^{2+} (Fig. 5a), GTP analogues (Fig. 3b) or PMA (data not shown)). Since both membrane fusion³⁶ and microtubule assembly³⁷ are sensitive to guanine nucleotides, it is possible that one of these could be the site at which control by guanine nucleotides is exerted in the exocytotic process.

We thank Professor F. Eckstein for gifts of purified guanine nucleotides and Professor C. A. Pasternak for supplies of Sendai virus. The program CHELATE was devised by Dr Sherwin Lee. The work was supported by grants from the Vandervell Trust and the MRC.

Received 4 April; accepted 11 December 1985.

- Nishizuka, Y. *Nature* **308**, 693-698 (1984).
- Rink, T. J. & Sanchez, A. *Biochem. J.* **222**, 833-836 (1984).
- Fernandez, J. M., Neher, E. & Gomperts, B. D. *Nature* **312**, 453-455 (1984).
- Gomperts, B. D. *Nature* **306**, 64-66 (1983).
- Cockcroft, S. & Gomperts, B. D. *Nature* **314**, 534-536 (1985).
- Cockcroft, S., Baldwin, J. M. & Allan, D. *Biochem. J.* **221**, 477-482 (1984).
- Wyke, A. M., Imprim, C. C., Knutton, S. & Pasternak, C. A. *Biochem. J.* **190**, 625-638 (1980).
- Bachi, T., Deas, J. E. & Howe, C. *Cell Surface Rev.* **2**, 83-127 (1977).
- Haywood, A. M. *J. molec. Biol.* **83**, 427-436 (1974).
- Foster, K. A., Gill, K., Micklem, C. A. & Pasternak, C. A. *Biochem. J.* **190**, 639-646 (1980).
- Imprim, C. C., Foster, K., Micklem, K. J. & Pasternak, C. A. *Biochem. J.* **186**, 847-860 (1980).
- Gomperts, B. D., Baldwin, J. M. & Micklem, K. J. *Biochem. J.* **210**, 737-745 (1983).
- Knight, D. E. & Baker, P. F. *J. Membrane Biol.* **68**, 107-140 (1982).
- Knight, D. E., Niggli, V. & Scrutton, M. C. *Eur. J. Biochem.* **143**, 437-446 (1984).
- Yaseen, M. A., Pedley, K. C. & Howell, S. L. *Biochem. J.* **206**, 81-87 (1982).
- Barrowman, M. M., Cockcroft, S. & Gomperts, B. D. *J. Physiol., Lond.* (submitted).
- Haslam, R. J. & Davidson, M. M. L. *FEBS. Lett.* **174**, 90-95 (1984).
- Baker, P. F. *Nature* **310**, 629-630 (1984).
- Cockcroft, S., Bennett, J. P. & Gomperts, B. D. *Biochem. J.* **200**, 501-508 (1981).
- Neidel, J. E., Kuhn, L. J. & Vandenberg, G. R. *Proc. natn. Acad. Sci. U.S.A.* **80**, 263-265 (1983).
- Cassel, D. & Pfeuffer, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2669-2673 (1978).
- Northup, J. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 6516-6520 (1980).
- Bokoch, G. M., Katada, T., Northup, J. K., Hewlett, E. L. & Gilman, A. G. *J. biol. Chem.* **258**, 2072-2075 (1983).
- Verghese, M. W., Smith, C. D. & Snydermann, R. *Biochem. biophys. Res. Commun.* **127**, 450-457 (1985).
- Smith, C. D., Lane, B. C., Kusaka, I., Verghese, M. W. & Snydermann, R. *J. biol. Chem.* **260**, 5875-5879 (1985).
- Becker, E. L. *et al. J. Cell Biol.* **100**, 1641-1646 (1985).
- Molski, T. F. P. *et al. Biochem. biophys. Res. Commun.* **124**, 644-650 (1984).
- Goldman, D. W., Chang, F.-H., Gifford, L. A., Goetzl, E. J. & Bourne, H. R. *J. exp. Med.* **162**, 145-156 (1985).
- Amsterdam, A., Ohad, I. & Schramm, M. J. *Cell Biol.* **41**, 753-773 (1969).
- Palade, G. *Science* **189**, 347-358 (1975).
- Chandler, D. E., Bennett, J. P. & Gomperts, B. D. *J. Ultrastruct. Res.* **83**, 221-232 (1983).
- Parsegian, V. A., Rand, R. P. & Gingell, D. *Ciba Fdn Symp.* **103**, 9-27 (1984).
- Palement, J. *Biochim. biophys. Acta* **777**, 274-282; *Expl Cell Res.* **151**, 354-366 (1984).
- Sandoval, I. V. & Weber, K. *J. biol. Chem.* **255**, 6966-6974 (1980).
- Bennett, J. P., Cockcroft, S. & Gomperts, B. D. *Biochim. biophys. Acta* **601**, 584-591 (1980).
- Andereg, G. *Helv. chim. Acta* **47**, 1801-1814 (1964).

Clonal origin of haematopoietic colonies in the postnatal mouse liver

J. Rossant*†, K. M. Vijn*†, C. E. Grossi† & M. D. Cooper†

* Department of Biological Sciences, Brock University, St Catharines, Ontario, Canada L2S 3A1

† Cellular Immunobiology Unit of the Tumor Institute and Department of Pathology, University of Alabama, Birmingham, Alabama 35294, USA

The liver of the neonatal mouse continues to show haematopoietic activity for up to 2 weeks after birth^{1,2} and morphological analysis has shown that this activity becomes focused in discrete haematopoietic colonies by the end of the first week postnatal³. Furthermore, each colony contains cells of one haematopoietic lineage only, that is, erythroid, myeloid or pre-B-lymphoid cells. This pattern of differentiation suggests that each colony is derived from a single committed precursor cell, which, if true, would represent the first demonstration of non-mixed haematopoietic colonies in normal development and would provide a useful system for studying the factors affecting the clonal diversity of haematopoietic stem cells and their lineage-committed progeny. Here we have analysed the haematopoietic foci in the liver of neonatal mouse chimaeras, using a newly developed ubiquitous *in situ* cell marker system which clearly demonstrates the clonal origin of these colonies.

Mouse chimaeras can be very useful for addressing problems of clonal development because they consist of a mixture of cells of two different genotypes, brought together by embryo aggregation or blastocyst injection at early stages of embryo development⁴. If a given structure in a chimaera is derived from a single progenitor cell, it will always be entirely of one genotype or the other, whereas structures with multiple progenitors will often

† Present address: Mt Sinai Hospital Research Institute, Division of Molecular and Developmental Biology, 600 University Avenue, Toronto, Ontario, Canada M5G 1X5.

- Douglas, W. W. *Br. J. Pharmacol.* **34**, 451-474 (1968).
- Sha'afi, R. I. *et al. Biochem. biophys. Res. Commun.* **114**, 638-645 (1983).
- Di Virgilio, F., Lew, D. P. & Pozzan, T. *Nature* **310**, 691-693 (1984).

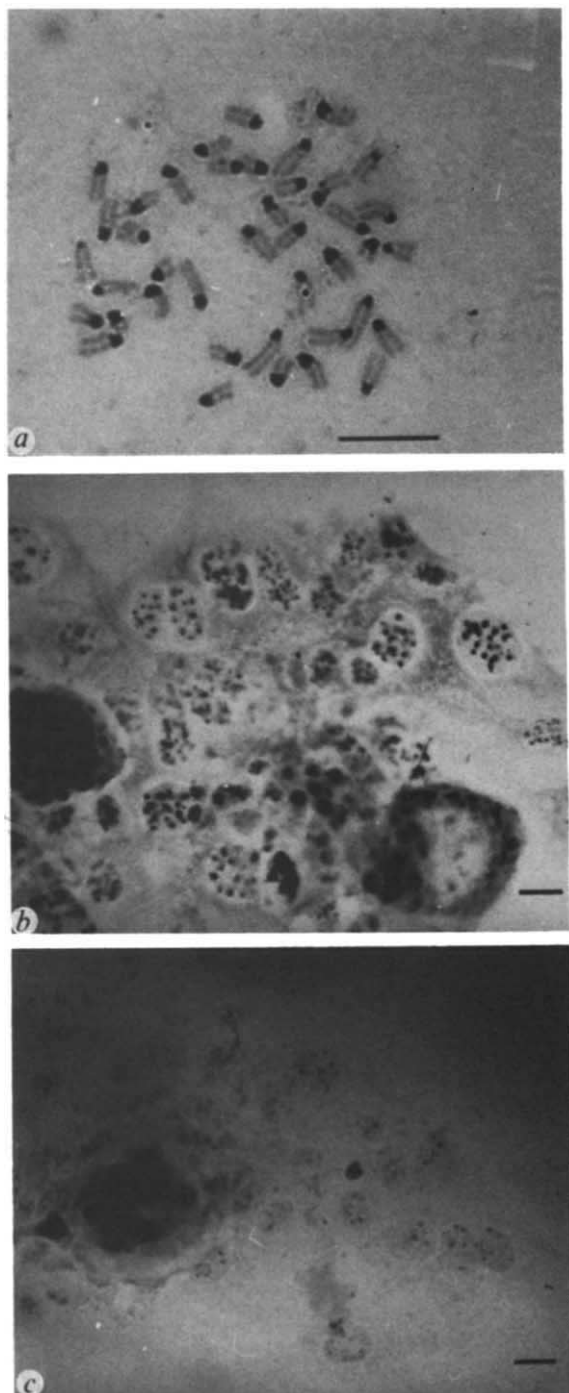


Fig. 1 Hybridization of biotinylated *M. musculus* satellite DNA to chromosomes and cell spreads. *a*, Chromosome spread from *M. musculus* showing hybridization of the probe to the centromeric regions alone. *b*, Blastocyst outgrowth from *M. musculus* showing extensive hybridization of the repetitive DNA probe as evidenced by dark clumps of staining in the nuclei. *c*, *M. caroli* blastocyst outgrowth showing only background staining. Scale bar, 20 μm .

Methods. The cloned probe to the *M. musculus* major satellite DNA sequence, pMR196¹⁰, was nick-translated in the presence of Bio-16-UTP (Enzo Biochem) according to the protocol provided by the manufacturer. Unincorporated nucleotides were separated from the labelled DNA by elution through a Sephadex G-50 spun column. Incorporation of biotinylated nucleotide into DNA was assessed by spotting serial dilutions of the nick-translated probe onto nitrocellulose and heating the latter at 80 °C for 1 h under vacuum. The filter was then incubated in 4% bovine serum albumin (BSA) in phosphate-buffered saline (PBS) for 10 min at room temperature, followed by incubation with streptavidin-horseradish peroxidase (HRP) complex (Detek I HRP, Enzo Biochem) as described by the manufacturers. After washing three times in PBS, the filter was stained in the dark for 2 min with 0.5 mg ml⁻¹ DAB (3,3'-diaminobenzidine tetrahydrochloride) in 0.05 M Tris-HCl pH 7.6, plus 0.01% hydrogen peroxide. Brown stain indicates incorporation of the biotinylated precursor and filters can be retained to allow comparison of incorporation from one batch to the next. Chromosome spreads were prepared from primary *M. musculus* embryonic fibroblast cultures in the standard manner. Blastocyst outgrowths were obtained by culturing blastocysts of the two species in chambered slides in α -modified minimal essential medium (MEM) plus 10% fetal calf serum for 4 days. Then, medium was removed and the cultures were fixed in acetic acid/ethanol (1:3) for 1 h, rinsed in ethanol and air-dried. Chromosome and cell spreads were pretreated for 1 h at 37 °C with 100 $\mu\text{g ml}^{-1}$ RNase A (Boehringer Mannheim) in 2 $\times$ SSC, washed twice in 2 $\times$ SSC, dehydrated and air-dried. Slides were denatured by immersing them in 70% formamide in 2 $\times$ SSC at 70 °C for 4 min, then they were dehydrated and air-dried. Biotinylated pMR196 (final concentration 3 $\mu\text{g ml}^{-1}$) was denatured by boiling in the presence of sonicated herring sperm DNA (final concentration 250 $\mu\text{g ml}^{-1}$) for 5 min. The probe mix was quenched on ice and 50% dextran sulphate and 20 $\times$ SSC were added to give a final concentration of 10% dextran sulphate and 4 $\times$ SSC. Probe mix (22 μl) was placed on each spread, which was covered by an acid-cleaned coverslip. Slides were incubated overnight at 60 °C in a moist chamber. Washes were carried out at fairly high stringency to avoid cross-hybridization. Two washes in 2 $\times$ SSC at room temperature were followed by two washes in 0.1 $\times$ SSC, first at room temperature and then at 50 °C. The slides were then treated with the streptavidin-HRP complex and the peroxidase activity was detected with DAB as described by the manufacturer. It is very important not to let the slides dry out at any point in this procedure. Slides were lightly counterstained with Light Green.

be of mixed genotype, depending on the proportion of the two genotypes in the chimaera and the degree of cell mixing during development^{4,5}. Determination of the genotype of a given structure requires sensitive genetic markers. The ideal cell marker should be ubiquitous, cell-autonomous and readily detectable in single cells in histological sections of chimaeras, so that the genotype of every cell in a structure can be determined unequivocally. There have been only a small number of *in situ* genetic marker systems described in mammals, and most have been limited to certain cell types or have not been readily detectable at the single-cell level⁶.

Recently, we have described the development of an *in situ* marker system for studying clonal development in mammalian chimaeras which overcomes many of the problems of the other

marker systems. Our system is based on the use of interspecific chimaeras between *Mus musculus*, the laboratory mouse, and *Mus caroli*, a wild species of mouse from South-East Asia⁷. *In situ* DNA-DNA hybridization with a cloned probe to the major satellite DNA sequence of *Mus musculus* distinguishes between the nuclei of the two species in both cell spreads and tissue sections^{8,9}, because *M. caroli* DNA contains very few sequences that cross-hybridize with the *M. musculus* satellite sequence¹⁰. This marker system has great potential for studying cell lineage in embryonic development and in the adult because it can be used in all tissues containing nucleated cells. In our original work we used a radioactively labelled probe combined with autoradiography to detect hybridization, but we have recently improved the reproducibility and resolution of the system by

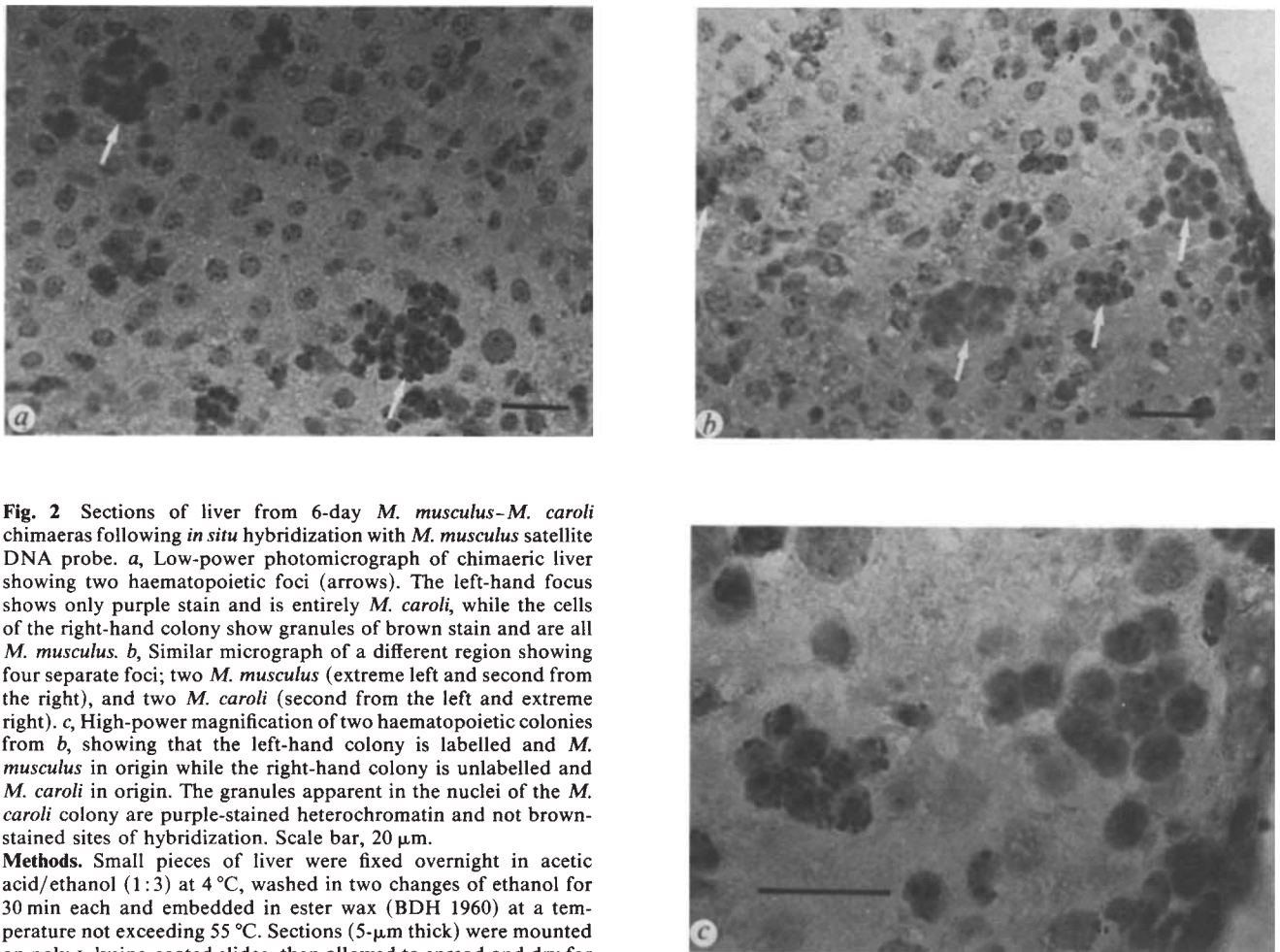


Fig. 2 Sections of liver from 6-day *M. musculus*-*M. caroli* chimaeras following *in situ* hybridization with *M. musculus* satellite DNA probe. *a*, Low-power photomicrograph of chimaeric liver showing two haematopoietic foci (arrows). The left-hand focus shows only purple stain and is entirely *M. caroli*, while the cells of the right-hand colony show granules of brown stain and are all *M. musculus*. *b*, Similar micrograph of a different region showing four separate foci; two *M. musculus* (extreme left and second from the right), and two *M. caroli* (second from the left and extreme right). *c*, High-power magnification of two haematopoietic colonies from *b*, showing that the left-hand colony is labelled and *M. musculus* in origin while the right-hand colony is unlabelled and *M. caroli* in origin. The granules apparent in the nuclei of the *M. caroli* colony are purple-stained heterochromatin and not brown-stained sites of hybridization. Scale bar, 20 μ m.

Methods. Small pieces of liver were fixed overnight in acetic acid/ethanol (1:3) at 4 °C, washed in two changes of ethanol for 30 min each and embedded in ester wax (BDH 1960) at a temperature not exceeding 55 °C. Sections (5- μ m thick) were mounted on poly-L-lysine-coated slides, then allowed to spread and dry for 1 h at 42 °C. Wax was removed using xylene and the slides were dipped in ethanol for 5 min before air-drying. Denaturation and hybridization were performed as described for the cell spreads in Fig. 1 legend. After staining for peroxidase the slides were run quickly through a haematoxylin and eosin staining series and mounted in Canada balsam. *M. musculus* nuclei show brown staining apparently associated with the heterochromatin, indicative of hybridization of the *M. musculus* satellite DNA probe. *M. caroli* nuclei stain purple with haematoxylin but show no brown staining. The difference between brown and purple stain is difficult to reproduce in a black-and-white photograph and so the distinction between labelled and unlabelled cells is not as striking in these figures as in the original slides. (Colour prints can be obtained from the authors on request.)

using biotinylated DNA¹¹ as the probe for hybridization (Fig. 1); this modification has transformed the marker system into one that can be used for the detailed single-cell analysis required to answer questions such as that addressed here.

Before analysing the genotype of haematopoietic foci in livers of neonatal chimaeras between *M. musculus* and *M. caroli*, we first examined the time course and nature of haematopoietic activity in *M. caroli*, to ensure that it was similar to that in *M. musculus*. Discrete haematopoietic colonies were observed slightly earlier in *M. caroli* than in *M. musculus*, but the cellular composition of the colonies in the two species was very similar (Table 1). Chimaeras between the two species were produced by injecting *M. caroli* inner cell masses into *M. musculus* blastocysts which were then transferred to *M. musculus* recipient females and allowed to proceed to term. Two chimaeras which contained approximately equal contributions from the two component species, by coat colour, were killed and their livers removed. Glucose phosphate isomerase isozyme analysis was performed on part of each liver¹² and the remainder was cut into small pieces and fixed and embedded for histological examination and *in situ* hybridization. The liver that was chosen for more extensive analysis was shown to contain ~70% *M. musculus* cells by isozyme analysis. Morphological analysis

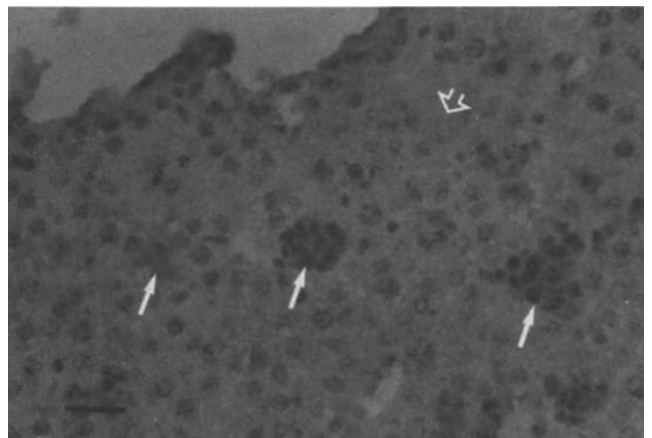


Fig. 3 Photomicrograph of chimaeric liver showing several haematopoietic foci, all of which are composed of cells of a single genotype (solid arrows). The small left-hand colony is *M. caroli* and the larger colonies are *M. musculus*. The open arrow indicates a megakaryocyte which is *M. caroli* in origin.

Table 1 Relative frequencies and cellular composition of liver haematopoietic foci as a function of day postnatal

Day postnatal	<i>M. musculus</i> (BALB/c)*				% Of colony type <i>M. caroli</i>				Interspecies chimaera			
	Erythroid	Myeloid	Lymphoid	Mega-karyocytes	Erythroid	Myeloid	Lymphoid	Mega-karyocytes	Erythroid	Myeloid	Lymphoid	Mega-karyocytes
6	Diffuse haematopoiesis				60	22	16	2	63	20	14	3
8	44	36	16	4	55	31	10	4		ND		
10	50†	34	5	10	66	26	8	0		ND		
12	77	16	4	3	Absent					ND		
15	Absent‡				Absent					ND		

Liver samples were fixed overnight in 10% formaldehyde, dehydrated and paraffin-embedded. Sections were stained with haematoxylin and eosin and colony types were identified on the basis of morphology. Other studies, using electron microscopy and immunofluorescence, have confirmed this morphological identification³. ND, not done.

* From ref. 3.

† 1% of these foci were mixed (erythroid plus myeloid).

‡ Foci no longer detectable.

Table 2 Distribution of *M. musculus* and *M. caroli* haematopoietic foci

<i>M. musculus</i>	71
<i>M. caroli</i>	19
Mixed type 1*	6
Mixed type 2†	4

Values are percentages.

* Colonies consisting of cells of the *M. caroli* type with a single cell of the *M. musculus* type.

† Colonies consisting of several cells from both species.

revealed that it contained discrete haematopoietic colonies whose cellular composition was very similar to those of non-chimaeric livers (Table 1). *In situ* hybridization with the *M. musculus* satellite DNA probe revealed roughly the expected proportion of *M. musculus* and *M. caroli* cells in the hepatocytes and also showed that there was fine intermixing of liver cells of the two genotypes, as found previously using the radioactively labelled probe on adult liver⁹ and also reported from studies using enzyme activity variants as cell markers¹³.

When the haematopoietic colonies in the chimaeric liver were examined for hybridization to the *M. musculus* satellite DNA probe, it was apparent that almost all colonies were entirely of one genotype or the other (Table 2, Figs 2, 3). The proportion of *M. musculus* to *M. caroli* colonies closely reflected the proportion of the two genotypes in the hepatocytes, illustrating again the high correlation coefficient between mosaic composition in different tissues of mouse chimaeras^{14,15}. Of 220 colonies scored, only 10% contained cells of both genotypes. There were, on average, 8 cells per colony in a single plane of section. Thus, if the cells present in the haematopoietic foci were the result of random distribution of non-clonally related cells, approximately 6% of the colonies might appear totally *M. musculus* by chance alone and 0.0006% of the colonies would appear totally *M. caroli*. The rest of the colonies would be expected to be mixed. The results clearly differ from this prediction. A high proportion of single-genotype colonies, as observed, would result if haematopoietic progenitors of the two genotypes were non-randomly distributed in the developing liver. This seems unlikely as colonies of both genotypes were distributed widely and evenly throughout the tissue. The remaining possibility is that *M. musculus* and *M. caroli* haematopoietic cells cannot coexist in close proximity within a haematopoietic focus. However, this seems highly unlikely in view of the evidence that hepatocytes and other cell types are intermixed in chimaeras¹⁶ and that cells of the two species interact normally even to the extent of forming chimaeric myotubes by cell fusion¹⁶. The most likely explanation of the large number of single-genotype foci is, therefore, that each colony contains the progeny of a single progenitor cell.

Some mixed colonies were observed; most of these consisted of a group of cells of one genotype surrounding one cell of the opposite genotype. We have shown previously that erythroid colonies in the neonatal liver often contain an accessory cell with macrophage-like features³ and it is tempting to speculate that the mixed colonies observed represent erythroid colonies in which erythroid cells and accessory cells happen to be of different genotype. Unfortunately, the treatment of the slides for *in situ* hybridization was not compatible with the detailed morphological examination required to confirm this directly. The remaining foci were truly mixed and probably represent the result of fusion of two neighbouring colonies. Occasional colonies of mixed lineage were also observed on morphological examination³ (Table 1).

The present study has clearly shown that the haematopoietic colonies which occur in the neonatal mouse liver are clonal in origin. As it has also been shown that each colony is derived from one haematopoietic lineage alone, it is probable that precursors committed to different lineages occur throughout the neonatal liver and give rise to discrete foci of cells. The alternative possibility that multipotential stem cells became restricted to different pathways by local environmental differences seems unlikely in the absence of any evidence for such microenvironmental variation in the liver. This represents the first demonstration of clonal development from single committed haematopoietic precursors in normal development and makes the neonatal liver system very useful for studying genetic and environmental factors that can affect the lineage preferences of haematopoietic stem cells. Our study also demonstrates the power of the *M. musculus*-*M. caroli* marker system, which provides a degree of resolution not afforded to date by any other marker system in mammals. The marker system can now be applied to many other problems of clonal analysis in mammalian chimaeras.

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (to J.R.) and by grants CA16673 and CA13148 from the National Cancer Institute of the NIH (to M.D.C.). J.R. is an E.W.R. Steacie Memorial Fellow of NSERC.

Received 10 June; accepted 26 November 1985.

1. Kamps, W. A. & Cooper, M. D. *J. Immun.* **129**, 526-531 (1982).
2. Velardi, A. & Cooper, M. D. *J. Immun.* **133**, 672-677 (1984).
3. Grossi, C. E., Velardi, A. & Cooper, M. D. *J. Immun.* **135**, 2303-2311 (1985).
4. McLaren, A. *Mammalian Chimaeras* (Cambridge University Press, 1976).
5. West, J. D. in *Development in Mammals* Vol. 3 (ed. Johnson, M. H.) 413-460 (Elsevier/North Holland, Amsterdam, 1978).
6. West, J. D. in *Chimaeras in Developmental Biology* (eds Le Douarin, N. & McLaren, A.) 39-67 (Academic, London, 1984).
7. Rossant, J. & Frels, W. I. *Science* **208**, 419-421 (1980).
8. Siracusa, L. D. *et al. J. Embryol. exp. Morph.* **73**, 163-178 (1983).

9. Rossant, J., Vijn, K. M., Siracusa, L. D. & Chapman, V. M. *J. Embryol. exp. Morph.* **73**, 179–191 (1983).
10. Pietras, D. F. *et al. Nucleic Acids Res.* **11**, 6965–6983 (1983).
11. Langer, P. R., Waldrop, A. A. & Ward, D. C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6633–6637 (1981).
12. Peterson, A. C., Frair, P. M. & Wong, G. G. *Biochem. Genet.* **16**, 681–690 (1978).
13. West, J. D. *J. Embryol. exp. Morph.* **36**, 151–161 (1976).
14. Falconer, D. S. & Avery, P. J. *J. Embryol. exp. Morph.* **43**, 195–219 (1978).
15. West, J. D., Bucher, T., Linke, I. M. & Dunnwald, M. *J. Embryol. exp. Morph.* **84**, 309–329 (1984).
16. Rossant, J. & Chapman, V. M. *J. Embryol. exp. Morph.* **73**, 193–205 (1983).

A human osteosarcoma cell line secretes a growth factor structurally related to a homodimer of PDGF A-chains

Carl-Henrik Heldin*, Ann Johnsson*,
Suzanne Wennergren*, Christer Wernstedt*,
Christer Betsholtz† & Bengt Westermark†

* Department of Medical and Physiological Chemistry, Box 575,
Biomedicum, S-751 23 Uppsala, Sweden

† Department of Pathology, University Hospital, S-751 85 Uppsala,
Sweden

Platelet-derived growth factor (PDGF), as purified from fresh human platelets, is a protein of relative molecular mass (M_r) 30,000 composed of two disulphide-linked subunit chains of similar size, named A and B (ref. 1). The dimer structure of PDGF seems to be important for its biological effects, as reduction irreversibly inactivates the factor; it is not known, however, whether PDGF exists as a heterodimer or as a mixture of homodimers. Amino-acid sequence analysis has revealed that the A- and B-chains of human PDGF are related to each other, and that the B-chain is almost identical to part of the *v-sis* gene product of simian sarcoma virus (SSV)^{2–5}. There is experimental evidence that a PDGF-like protein is indeed operational in SSV-induced transformation^{6–12} and the biologically active *v-sis* product is probably structurally similar to a putative dimer of PDGF B-chains¹³. PDGF-like growth factors and/or a 4.2-kilobase (kb) *c-sis* transcript are present in several transformed mammalian cell lines^{7,14–24} and in certain nontransformed cells^{25–30}; cloned *c-sis* complementary DNA from human T cells transformed with human T-lymphotropic virus (HTLV)^{24,31} or from human endothelial cells³² contains the coding sequence for a putative PDGF B-chain precursor, but apparently lacks PDGF A-chain sequences. We have previously partially purified and characterized a PDGF-like growth factor from U-2 OS cells (osteosarcoma-derived growth factor, ODGF) and shown that this factor has structural, functional and immunological characteristics in common with PDGF¹⁴. We describe here a procedure for the preparation of homogeneous ODGF, and provide evidence that this factor, which binds to the PDGF receptor, has a structure similar to a homodimer of PDGF A-chains.

Serum-free conditioned medium was collected from U-2 OS cells³³ grown in roller bottles and used as starting material for the purification of ODGF. The growth factor activity was monitored by a ³H-thymidine incorporation assay using human foreskin fibroblasts as target cells³⁴, and occasionally by a PDGF receptor competition assay³⁵. ODGF was first concentrated and freed of glycosaminoglycans and anionic proteins by adsorption to a highly charged Sulphadex gel. The sample was then subjected to two consecutive gel chromatographies, the first at high ionic strength and neutral pH (Fig. 1a), the second at low pH (Fig. 1b). Final purification was achieved by reverse-phase chromatography on HPLC (Fig. 1c). This purification procedure gave a 4,000-fold purification at a yield of about 30% relative to the starting conditioned medium. Approximately 25 µg of ODGF was obtained from 8 l of conditioned medium. The

purified factor had PDGF agonist activity and was recognized by antibodies against PDGF (not shown).

Purified ODGF was analysed by SDS-gel electrophoresis and silver staining, and one homogeneous component of M_r 31,000 was obtained in non-reducing conditions (Fig. 1d). The size of this component corresponds to the largest PDGF component, which is proteolysed in the platelets or during purification to several components of M_r s 31,000–28,000 (refs 36–40 and Fig. 1d). In reducing conditions ODGF appeared as a single band of M_r 17,000, whereas PDGF contained several components of M_r 17,000 and smaller (Fig. 1d).

The behaviour of ODGF in SDS-gel electrophoresis in the absence and presence of reducing agents indicates that, like PDGF, it has a dimer structure. Furthermore, the homogeneous appearance of ODGF after reduction suggests that it may contain only one of the two prototype chains of PDGF. To investigate this possibility, ODGF and PDGF were radioactively iodinated using two different techniques, the chloramine-T method⁴¹ which labels tyrosine residues, and the procedure of Bolton and Hunter⁴² which labels amino groups (lysine residues and the amino terminus). The structure of the labelled component was investigated by SDS-gel electrophoresis and autoradiography. Both labelling procedures led to incorporation of ¹²⁵I into a 31,000- M_r component of ODGF which was converted to a 17,000- M_r form after reduction (Fig. 2), that is, the same pattern as was obtained with silver staining. PDGF, however, gave different patterns by the two labelling procedures. The chloramine-T method led to the labelling of components of apparent M_r 31,000–28,000, which were converted to apparent M_r s of 17,000, 16,000 and 12,000 after reduction. The Bolton-Hunter procedure led to incorporation of radioactivity into the same components, but also into a component of apparent M_r 15,000 and into two components of M_r s <10,000 after reduction (Fig. 2). The different patterns obtained by the two labelling procedures are explained by the fact that the PDGF B-chain contains no tyrosine residues⁵ and is therefore not labelled by the chloramine-T method. This conclusion was supported by experiments in which the constituent chains of ¹²⁵I-PDGF were separated by reverse-phase chromatography (see below); labelling of PDGF by the chloramine-T method yielded radioactivity only in the position corresponding to the PDGF A-chain (not shown). It follows from the labelling pattern of ODGF that the factor may contain the PDGF A-chain (Fig. 2).

Furthermore, reduced and alkylated ¹²⁵I-ODGF, labelled using the Bolton-Hunter procedure, was subjected to a reverse-phase chromatographic system capable of separating the A- and B-chains of PDGF³⁶. Only one peak of radioactivity was obtained, with an elution position corresponding exactly to that of PDGF A-chains (Fig. 3a, b).

In addition to their different elution positions in reverse-phase chromatography, the A- and B-chains of PDGF differed in their susceptibilities to cleavage by formic acid and cyanogen bromide. Purified ¹²⁵I-labelled PDGF A-chain was cleaved by acid treatment (70% formic acid for 48 h at 37 °C) from an M_r of 17,000 to one of 15,000, while this treatment had no effect on the B-chain (Fig. 3c). This cleavage probably occurs at an aspartic acid-proline linkage at position 25–26 in the A-chain, since such a peptide bond is known to be susceptible to acidic conditions⁴³; no such peptide bond occurs in the PDGF B-chain⁵. However, the B-chain was cleaved by CNBr in 70% formic acid, leading to the formation of two new peptides of M_r s 14,000 and 12,000 (Fig. 3c). The cleavages may occur at a methionine residue at position 12 and at a tryptophan residue at position 40 in the B-chain². Incubation of the A-chain with CNBr in 70% formic acid produced no observable cleavage in addition to that caused by 70% formic acid only (Fig. 3c). This is compatible with the fact that the known sequence of the A-chain contains no methionine residues and that the only tryptophan residue is at position 36, close to the aspartic acid-proline linkage⁵ at position 25–26. When reduced and alkylated

125 I-ODGF was subjected to acid treatment and CNBr cleavage, practically all of it was cleaved by acid to an M_r of 15,000, but CNBr had no further effect (Fig. 3c), that is, it behaved exactly as if it consisted entirely of PDGF A-chains.

Amino-acid sequence analysis of the N-terminus of ODGF further substantiated the presence of only A-chains in this molecule. About 200 pmol of unreduced ODGF was analysed using a solid-phase sequencer, yielding an 8-amino-acid N-terminal sequence. This contained only one major sequence, which was identical to the N-terminal sequence of the A-chain of PDGF (Fig. 4). When unreduced PDGF purified from human

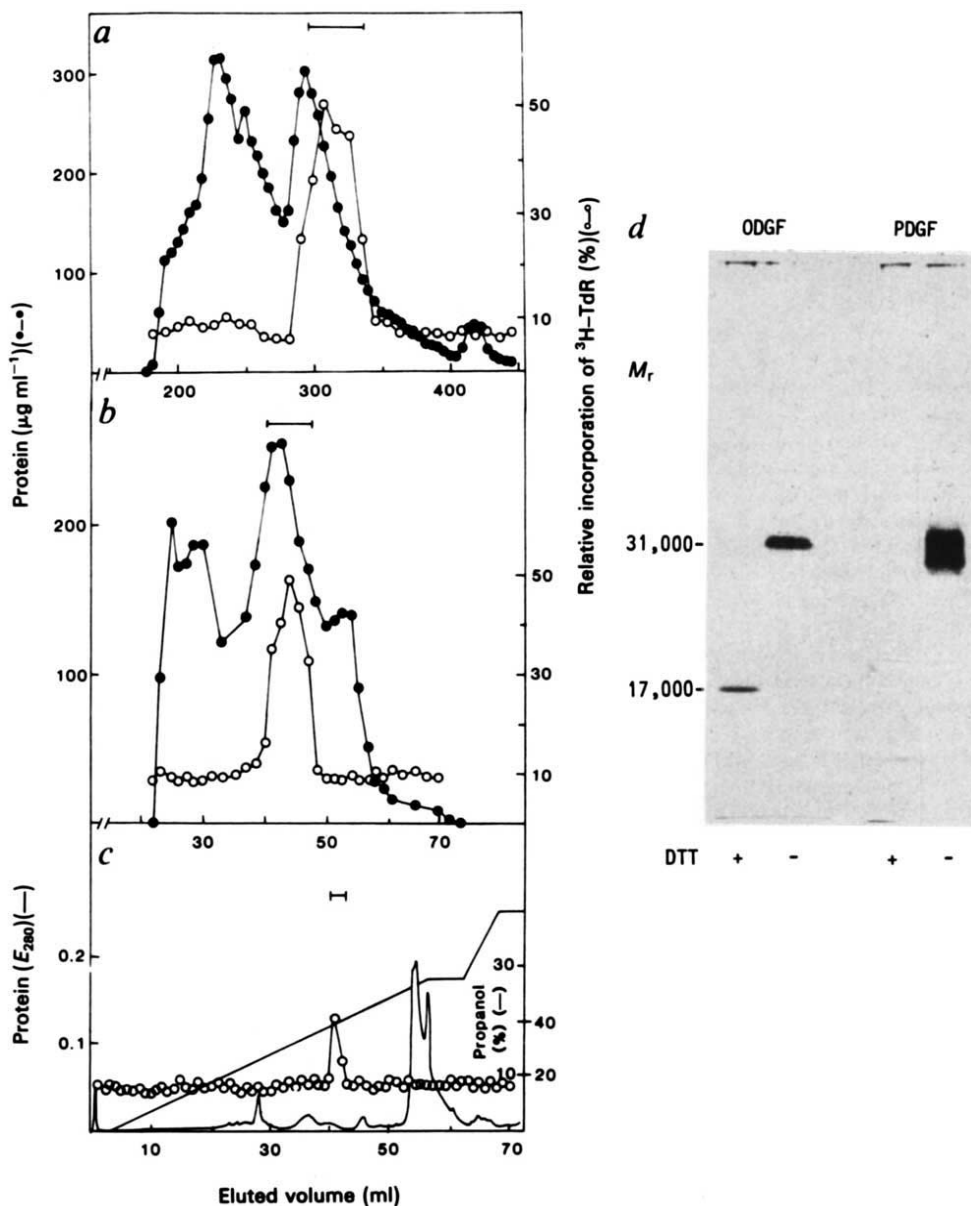
platelets was applied to the same analysis, sequences from both the A- and B-chains were obtained in approximately stoichiometric quantities (not shown).

We report here, for the first time, the purification to homogeneity of a PDGF-like growth factor from a human transformed cell line, the U-2 OS osteosarcoma cell line. The factor appears to be a homodimer of polypeptide chains similar to the PDGF A-chain. This is unexpected, because previous investigations have shown that the U-2 OS cells in fact express the *c-sis* gene, which encodes the B-chain precursor; a 4.2-kb messenger RNA has been identified in Northern blots using a *c-sis* probe

Fig. 1 Purification of ODGF.

Sequential chromatography of conditioned medium from U-2 OS cells on a Sephacryl S-200 column eluted with 1 M NaCl, 0.01 M phosphate buffer pH 7.4 (a), a BioGel P-150 column eluted with 1 M acetic acid (b), and an HPLC RP8 column (c), are shown, as well as a silver-stained SDS gel of the purified ODGF (d).

Methods. Conditioned medium from U-2 OS cells was collected as described previously¹⁴. The first step in the purification of ODGF was batch-wise adsorption to Sulphadex beads. Sulphadex gel (200 ml, synthesized as described elsewhere⁴⁸) was added to 8 l of conditioned medium (containing ~320 mg protein) and then equilibrated overnight at 4 °C on a rocker platform. The gel was then allowed to sediment, the unadsorbed fraction discarded and the gel poured into a column (5 × 40 cm). The column was washed with 600 ml of 0.15 M NaCl, 0.01 M phosphate buffer pH 7.4 and then eluted with 300 ml of 1.5 M NaCl, 0.01 M phosphate buffer pH 7.4. The eluted fraction contained ~60 mg of protein. ODGF was then concentrated by addition of solid ammonium sulphate to 60% saturation (390 g l⁻¹). After 2 h of equilibration end over end at 4 °C, the sample was centrifuged for 20 min at 20,000 g. The pellet was dissolved in 5 ml of 1 M NaCl, 0.01 M phosphate buffer pH 7.4 and then dialysed against this buffer. After centrifugation the sample was applied to a Sephacryl S-200 column (2 × 145 cm) and eluted at 4 °C with 1 M NaCl, 0.01 M phosphate buffer pH 7.4 at a flow rate of 10 ml h⁻¹. Fractions (4 ml) were collected and analysed for protein and growth promoting activity (stimulation of ³H-thymidine incorporation in human foreskin fibroblasts³⁴). Growth promoting activity is expressed relative to that given by 10 ng ml⁻¹ of PDGF (5–10 × 10³ c.p.m., corresponding to 50–70% labelled nuclei). Active fractions were pooled as indicated in the figure (~8 mg of protein), dialysed against 1 M acetic acid and lyophilized. The sample was dissolved in 2 ml of 1 M acetic acid and applied to a BioGel P-150 column (2 × 30 cm) then eluted with 1 M acetic acid. The column was operated at room temperature, at a flow rate of 4 ml h⁻¹ and 1.5-ml fractions were collected. The active fractions were pooled as indicated in the figure (~1 mg of protein) and lyophilized. The lyophilized material was then dissolved in 300 µl of 1 M acetic acid and applied to an HPLC RP8 column (Merck). The column was washed with 4 ml of 1 M acetic acid, 2 M guanidine-HCl, and then eluted with a linear gradient of propanol in 1 M acetic acid, 2 M guanidine-HCl. Fractions containing growth-promoting activity were pooled as indicated in the figure and used for further experiments. About 25 µg of ODGF was obtained from 8 l of conditioned medium. Aliquots (~0.5 µg) of ODGF, and PDGF purified from human platelets³⁶, were analysed on gradient gels consisting of 13–18% acrylamide, in non-reducing conditions or after reduction with dithiothreitol (DDT) and alkylation⁴⁹. The gel was stained for protein by a silver method⁵⁰. Note that ODGF and PDGF stain better in unreduced than reduced form. The protein concentration of purified ODGF was estimated by comparing the intensity of the ODGF band in the silver-stained SDS gel with those of known amounts of PDGF.



Active fractions were pooled as indicated in the figure (~8 mg of protein), dialysed against 1 M acetic acid and lyophilized. The sample was dissolved in 2 ml of 1 M acetic acid and applied to a BioGel P-150 column (2 × 30 cm) then eluted with 1 M acetic acid. The column was operated at room temperature, at a flow rate of 4 ml h⁻¹ and 1.5-ml fractions were collected. The active fractions were pooled as indicated in the figure (~1 mg of protein) and lyophilized. The lyophilized material was then dissolved in 300 µl of 1 M acetic acid and applied to an HPLC RP8 column (Merck). The column was washed with 4 ml of 1 M acetic acid, 2 M guanidine-HCl, and then eluted with a linear gradient of propanol in 1 M acetic acid, 2 M guanidine-HCl. Fractions containing growth-promoting activity were pooled as indicated in the figure and used for further experiments. About 25 µg of ODGF was obtained from 8 l of conditioned medium. Aliquots (~0.5 µg) of ODGF, and PDGF purified from human platelets³⁶, were analysed on gradient gels consisting of 13–18% acrylamide, in non-reducing conditions or after reduction with dithiothreitol (DDT) and alkylation⁴⁹. The gel was stained for protein by a silver method⁵⁰. Note that ODGF and PDGF stain better in unreduced than reduced form. The protein concentration of purified ODGF was estimated by comparing the intensity of the ODGF band in the silver-stained SDS gel with those of known amounts of PDGF.

in stringent hybridization conditions^{44,45}. If PDGF B-chains are indeed also synthesized in the U-2 OS cells, there are several possible reasons that we have not observed any such peptides in the ODGF preparation; for example, B-chains may not be properly dimerized, processed or secreted into the culture medium. Alternatively, B-chain homodimers or A- and B-chain heterodimers are formed and secreted, but are turned over rapidly, or remain associated with the cells^{12,46}, and are therefore not found in the conditioned medium. It is also possible that B-chain-containing molecules appear in the conditioned medium but are lost during purification.

PDGF purified from human platelets contains approximately equal quantities of A- and B-chains (ref. 5 and our data not shown). Fractionation of purified PDGF by various methods has not yielded any evidence for the presence of homodimers, and preliminary binding experiments indicate that PDGF from human platelets interacts with the PDGF receptor at higher affinity than ODGF (not shown). Taken together, these observations suggest that human PDGF may have a heterodimer structure consisting of both A- and B-chains which interacts with the PDGF receptor more avidly than the osteosarcoma-derived A-chain homodimer. However, the extensive proteolysis of

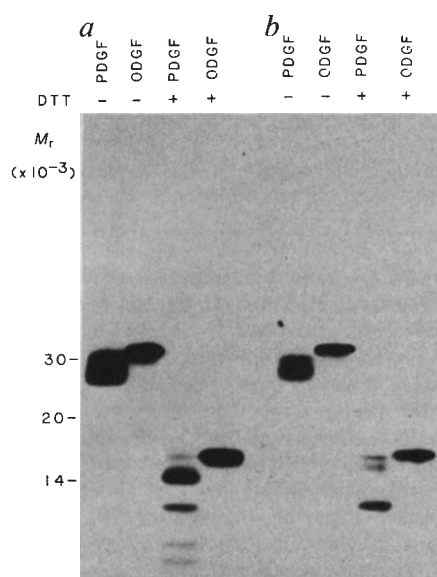


Fig. 2 Analysis by SDS-gel electrophoresis and autoradiography of ODGF and PDGF labelled with ¹²⁵I by two different methods: *a*, according to the method of Bolton and Hunter⁴²; *b*, by the chloramine-T method⁴¹. Samples were analysed with or without reduction with DDT and alkylation. ODGF (1 µg) and PDGF (5 µg) were labelled by the procedure of Bolton and Hunter⁴² to specific activities of 42,000 c.p.m. per ng and 32,000 c.p.m. per ng, respectively. Furthermore, 1 µg of ODGF and 5 µg of PDGF were labelled with the chloramine-T method⁴¹, as described previously³⁵, to specific activities of 90,000 and 50,000 c.p.m. per ng, respectively.

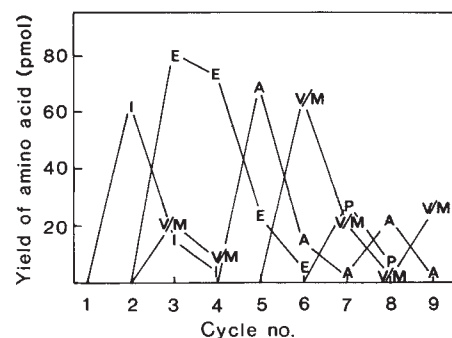
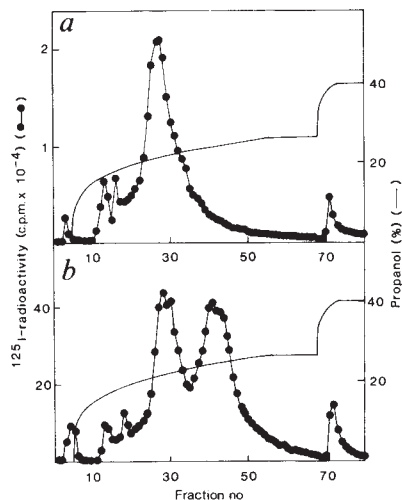
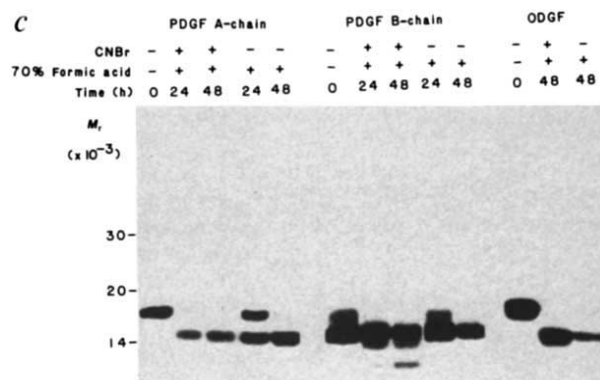


Fig. 4 Analysis of the amino-acid sequence of non-reduced ODGF. The yield of the PTH amino acids released at each step of Edman degradation is plotted against the cycle number. The solid-phase technique was used⁵¹. ODGF was coupled to diisothiocyanate-treated glass beads of pore size 120 Å. Data are shown from one run using ~200 pmol of ODGF. The assigned N-terminal amino-acid sequence of ODGF is compared with those of the A- and B-chains of PDGF. X denotes unidentified amino acid. V/M denotes valine or methionine (the PTH derivatives of these amino acids were not separated in the HPLC system used for identification).

Fig. 3 Separation of the constituent polypeptide chains of ¹²⁵I-labelled ODGF (*a*) and ¹²⁵I-labelled PDGF (*b*). ¹²⁵I-ODGF and ¹²⁵I-PDGF (both labelled by the Bolton-Hunter procedure⁴²) were reduced and alkylated and applied to an HPLC RP8 column (Merck). The column was eluted in 1 M acetic acid, 2 M guanidine-HCl with a nonlinear gradient of propanol, as described previously³⁶. *c*, Gel showing the susceptibility of the constituent chains of PDGF and ODGF to chemical fragmentation.



Separated ¹²⁵I-labelled A- and B-chains of PDGF, and reduced and alkylated ¹²⁵I-ODGF, were incubated with 70% formic acid or 30 mg ml⁻¹ of CNBr in 70% formic acid for 24 or 48 h at 37 °C. Samples were then analysed by SDS-gel electrophoresis on 13–18% acrylamide gradient gels and subjected to autoradiography. The 17,000-*M_r* band in the B-chain preparation represents a slight contamination of A-chains.



PDGF purified from human platelets makes it difficult to form any definite conclusions about its subunit structure. Interestingly, pig PDGF seems to be a homodimer of polypeptide chains similar to the human B-chain and thus similar in structure to the PDGF-like factor from SSV-transformed cells (ref. 47 and M. Waterfield, personal communication).

The present work raises some interesting questions about the regulation of the expression of the PDGF genes. The A- and B-chains are probably encoded by different genes, and thus translated by different transcripts. Apparently, either chain can be assembled into a homodimer structure after translation, A-A in the case of ODGF and B-B in the cases of pig PDGF and the v-sis gene product. Activation of either PDGF gene may lead to the production of PDGF agonist activity, as both types of homodimers interact with the PDGF receptor. If this occurs in a cell that carries PDGF receptors, an autocrine loop may be formed, leading to stimulation of cell growth, as discussed for U-2 OS cells⁴⁴ and found to be the case in SSV-transformed cells¹². It will be of particular interest to determine whether human PDGF and the PDGF-like factors produced in a regulated manner by normal cells are homodimers or heterodimers, and to compare the functional activities of the different but homologous PDGF-like growth factors from normal and malignant cells.

We thank Ulf Hellman for advice on amino-acid sequencing and Marianne Kastemar for help with cell culturing. These studies were supported by grants from the Swedish Cancer Society (689, 786 and 1794), Konung Gustaf V:s 80-årsfond and the Swedish Department of Agriculture.

Received 21 October; accepted 17 December 1985.

1. Heldin, C.-H., Wasteson, Å. & Westermark, B. *Molec. cell. Endocr.* **39**, 169-187 (1985).
2. Waterfield, M. D. *et al. Nature* **304**, 35-39 (1983).
3. Doolittle, R. F. *et al. Science* **221**, 275-277 (1983).
4. Devare, S. G., Reddy, E. P., Law, J. D., Robbins, K. C. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 731-735 (1983).
5. Johnsson, A. *et al. EMBO J.* **3**, 921-928 (1984).
6. Deuel, T. F., Huang, J. S., Huang, S. S., Stroobant, P. & Waterfield, M. D. *Science* **221**, 1348-1350 (1983).
7. Bowen-Pope, D. F., Vogel, A. & Ross, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2396-2400 (1984).
8. Owen, A. J., Pantazis, P. & Antoniades, H. N. *Science* **225**, 54-56 (1984).
9. Garrett, J. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 7466-7470 (1984).
10. Huang, J. S., Huang, S. S. & Deuel, T. F. *Cell* **39**, 79-87 (1984).
11. Johnsson, A., Betsholtz, C., von der Helm, K., Heldin, C.-H. & Westermark, B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1721-1725 (1985).
12. Johnsson, A., Betsholtz, C., Heldin, C.-H. & Westermark, B. *Nature* **317**, 438-440 (1985).
13. Robbins, K. C., Antoniades, H. N., Devare, S. G., Hunkapiller, M. W. & Aaronson, S. A. *Nature* **305**, 605-608 (1983).
14. Heldin, C.-H., Westermark, B. & Wasteson, Å. *J. cell. Physiol.* **105**, 235-246 (1980).
15. Graves, D. T., Owen, A. J., Antoniades, H. N. *Cancer Res.* **43**, 83-87 (1983).
16. Betsholtz, C. *et al. Biochem. biophys. Res. Commun.* **117**, 176-182 (1983).
17. Nistér, M., Heldin, C.-H., Wasteson, Å. & Westermark, B. *Proc. natn. Acad. Sci. U.S.A.* **81**, 926-930 (1984).
18. Stroobant, P., Gullick, W. J., Waterfield, M. D. & Rozengurt, E. *EMBO J.* **4**, 1945-1949 (1985).
19. Betsholtz, C., Bywater, M., Westermark, B., Bürk, R. R. & Heldin, C.-H. *Biochem. biophys. Res. Commun.* **130**, 753-760 (1985).
20. Gudas, L. J., Singh, J. P. & Stiles, C. D. *Cold Spring Harb. Conf. Cell Proliferation* **10**, 229-236 (1983).
21. Van Zoelen, E. J. *et al. Molec. cell. Biol.* **5**, 2289-2297 (1985).
22. Eva, A. *et al. Nature* **295**, 116-119 (1982).
23. Norris, J. S. *et al. Biochem. biophys. Res. Commun.* **122**, 124-128 (1984).
24. Josephs, S. F., Guo, C., Ratner, L. & Wong-Staal, F. *Science* **223**, 487-490 (1984).
25. DiCorleto, P. E. & Bowen-Pope, D. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1919-1923 (1983).
26. Barrett, T. B., Gajdusek, C. M., Schwartz, S. M., McDougall, J. K. & Benditt, E. P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6772-6774 (1984).
27. Jaye, M. *et al. Science* **228**, 882-885 (1985).
28. Seifert, R. A., Schwartz, S. M. & Bowen-Pope, D. F. *Nature* **34**, 669-671 (1984).
29. Nilsson, J., Sjöberg, M., Palmberg, L., Thyberg, J. & Heldin, C.-H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4418-4422 (1985).
30. Goustin, A. S. *et al. Cell* **41**, 301-312 (1985).
31. Ratner, L. *et al. Nucleic Acids Res.* **13**, 5007-5018 (1985).
32. Collins, T., Ginsburg, D., Boss, J. M., Orkin, S. H. & Pober, J. S. *Nature* **316**, 748-750 (1985).
33. Pontén, J. & Saksela, E. *Int. J. Cancer* **2**, 434-447 (1967).
34. Betsholtz, C. & Westermark, B. *J. cell. Physiol.* **118**, 203-210 (1984).
35. Heldin, C.-H., Westermark, B. & Wasteson, Å. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3664-3668 (1981).
36. Johnsson, A., Heldin, C.-H., Westermark, B. & Wasteson, Å. *Biochem. biophys. Res. Commun.* **104**, 66-74 (1982).
37. Heldin, C.-H., Westermark, B. & Wasteson, Å. *Biochem. J.* **193**, 907-913 (1981).
38. Antoniades, H. N. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7314-7317 (1981).
39. Deuel, T. F. *et al. J. Biol. Chem.* **256**, 8896-8899 (1981).
40. Raines, E. W. & Ross, R. *J. Biol. Chem.* **257**, 5154-5160 (1982).
41. Hunter, W. M. & Greenwood, F. C. *Nature* **194**, 495-496 (1962).
42. Bolton, A. E. & Hunter, W. M. *Biochem. J.* **133**, 529-539 (1973).
43. Piszczewicz, D., Landon, M. & Smith, E. L. *Biochem. biophys. Res. Commun.* **40**, 1173-1178 (1970).
44. Betsholtz, C., Westermark, B., Ek, B. & Heldin, C.-H. *Cell* **39**, 447-457 (1984).
45. Graves, D. T. *et al. Science* **226**, 972-974 (1984).
46. Robbins, K. C., Leal, F., Pierce, J. H. & Aaronson, S. A. *EMBO J.* **4**, 1783-1792 (1985).
47. Stroobant, P. & Waterfield, M. D. *EMBO J.* **3**, 2963-2967 (1984).
48. Miletic, J. P., Broze, G. J. Jr & Majerus, P. W. *Analyt. Biochem.* **105**, 304-310 (1980).
49. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835-851 (1975).
50. Morrissey, J. H. *Analyt. Biochem.* **117**, 307-310 (1981).
51. Laurssen, R. A. & Machleidt, W. *Meth. biochem. Analysis* **26**, 201-284 (1980).

A saturable receptor for ³²P-inositol-1,4,5-trisphosphate in hepatocytes and neutrophils

András Spät*, Peter G. Bradford, Jerry S. McKinney, Ronald P. Rubin & James W. Putney Jr

Division of Cellular Pharmacology, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298, USA

Several receptors for neurotransmitters, hormones and growth factors cause accelerated phosphodiesteratic breakdown of polyphosphoinositides when activated^{1,2}. One of the soluble products of this reaction, inositol-1,4,5-trisphosphate (Ins(1,4,5)P₃) is thought to act as a second messenger signalling the release of Ca²⁺ from intracellular stores³. In support of this hypothesis, several studies have shown that Ins(1,4,5)P₃ releases sequestered Ca²⁺ from permeable cells⁴⁻⁶ and microsomes⁷⁻⁹. On the basis of certain structural requirements for Ca²⁺-releasing activity by inositol phosphates, it has been postulated that Ins(1,4,5)P₃ acts by binding to a specific intracellular receptor, probably on a component of the endoplasmic reticulum¹⁰. Here we report that ³²P-Ins(1,4,5)P₃ binds to a specific saturable site in permeabilized guinea pig hepatocytes and rabbit neutrophils, and that the properties of this binding site suggest that it is the physiological receptor for Ins(1,4,5)P₃.

High-specific-activity ³²P-Ins(1,4,5)P₃ (~40 Ci mmol⁻¹) was prepared by phosphorylation of red cell ghosts with ³²P-ATP (3,000 Ci mmol⁻¹), followed by hydrolysis and purification as described in Fig. 1 legend. Binding experiments were carried out on ice and in the presence of 1 or 2 mM diphosphoglycerate to inhibit any metabolism of the ligand by, or binding of the ligand to, inositol-1,4,5-trisphosphate 5-phosphomonoesterase¹¹. In a few experiments, increasing the concentration of diphosphoglycerate to 10 mM did not affect total binding, indicating that the binding site was not the 5-phosphomonoesterase. We used guinea pig hepatocytes and rabbit neutrophils whose plasma membranes were made permeable with saponin as described previously¹². We used permeable cells rather than membranes because binding could be more readily demonstrated and correlated with Ca²⁺ release experiments performed on the same preparation. Because binding and dissociation were rapid (Fig. 1), a non-diluting technique was used to separate bound from free ligand. The technique involved centrifugation of suspensions of permeable cells through silicon oil (see Fig. 1 legend). The cells (and some trapped volume which can be corrected for with an inert marker such as ³H-mannose) sediment at the bottom of the tubes, whereas the medium containing free ligand remains above the oil. No dilution of the sample is necessary, and the technique should be generally applicable to the study of binding of any ligand which dissociates rapidly.

Association of ³²P-Ins(1,4,5)P₃ with its specific binding site in neutrophils is rapid, even at 0-5 °C, and binding is readily reversible (Fig. 1). The binding of ³²P-Ins(1,4,5)P₃ to hepatocytes

* Present address: Department of Physiology, Semmelweis University Medical School, Budapest, Hungary.

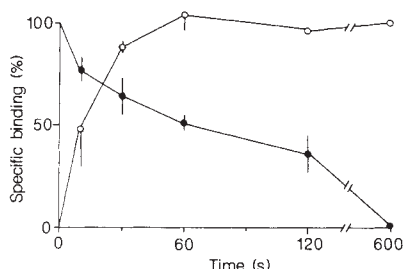


Fig. 1 Time course of: ○, binding of ^{32}P -Ins(1,4,5) P_3 to, and ●, dissociation of bound ^{32}P -Ins(1,4,5) P_3 from, saponin-permeabilized rabbit neutrophils. For the dissociation experiments $10\text{ }\mu\text{M}$ Ins(1,4,5) P_3 was added to the incubation mixtures ($t=0$) after incubation for 10 min in the presence of ^{32}P -Ins(1,4,5) P_3 and the time course of decrease in specific binding was determined. Binding data for both curves are expressed as a percentage of specific binding at $t=10$ min (just before the addition of non-radioactive Ins(1,4,5) P_3). Values shown are mean $\pm$ s.e.m. of 3–6 separate experiments, except for the association curve at $t=120$ s, which is the mean of two.

Methods. Rabbit neutrophils¹⁵ and guinea pig hepatocytes¹² were prepared as described previously, permeabilized with saponin and incubated in a high- K^+ salt solution as described previously¹², with $[\text{Ca}^{2+}]$ buffered with EGTA to 180 nM . Binding was determined by incubating permeable cells with ligand in this intracellular type medium and then at designated intervals, centrifuging the sample with a microfuge through $500\text{ }\mu\text{l}$ silicon oil (Nyosil; Nye, Inc., New Bedford, Massachusetts). Typical incubation conditions for neutrophils were $300\text{--}500\text{ }\mu\text{g}$ cellular protein and $0.02\text{ }\mu\text{Ci}$ ^{32}P -Ins(1,4,5) P_3 in $500\text{ }\mu\text{l}$, and for hepatocytes, $1\text{--}2\text{ mg}$ cellular protein and $0.05\text{ }\mu\text{Ci}$ ^{32}P -Ins(1,4,5) P_3 in 1.0 ml . For all samples, ^3H -mannose was included to allow correction for trapped ^{32}P in the silicon oil pellets. Nonspecific binding was determined by including $10\text{ }\mu\text{M}$ Ins(1,4,5) P_3 in the incubation medium. Specific binding generally exceeded 100 d.p.m. All samples were counted for a sufficient interval to ensure $>90\%$ reliability in radioactivity estimates. Preparation of ^{32}P -Ins(1,4,5) P_3 . Human erythrocyte ghosts were incubated for 20 min at room temperature in a solution modified from ref. 16 containing (per ml): $5\text{ }\mu\text{mol}$ MgCl_2 , $1\text{ }\mu\text{mol}$ EGTA, $25\text{ }\mu\text{mol}$ imidazole-HCl (pH 7.0), 20 mCi ^{32}P -ATP ($>3,000\text{ Ci mmol}^{-1}$; NEN). Phosphorylation was stopped by adding 8 vol. of 1 mM EDTA in 10 mM Tris buffer (pH 7.5). Labelled polyphosphoinositides were hydrolysed for 10 min at room temperature in 0.18 mM CaCl_2 in 25 mM imidazole buffer (pH 7.0). The reaction was stopped with 10% perchloric acid, neutralized with KOH, and buffered to pH 7.5–8.5 with 50 mM Tris-phosphate buffer. ^{32}P -Ins(1,4,5) P_3 was separated from other soluble radioactivity by anion-exchange chromatography as described by Berridge *et al.*¹⁷. Ammonium formate in the inositol triphosphate fraction was converted to formic acid by passing over Dowex-50 (H^+). The eluate was lyophilized and stored at -76°C . Radiochemical purity was $>90\%$ as assessed by paper chromatography¹⁸ and HPLC¹⁹. Specific activity was estimated by displacement of specific binding and was $30\text{--}50\text{ Ci mmol}^{-1}$.

(data not shown) was as rapid, and specific binding at 30 s was not distinguishable from that at 10 min. Nonspecific binding was taken as the residual binding in the presence of $10\text{ }\mu\text{M}$ Ins(1,4,5) P_3 ($30\text{ }\mu\text{M}$ Ins(1,4,5) P_3 did not reduce binding further). After correction for free ligand in the trapped volume of the pellet, nonspecific binding was 15–20% of the total for the neutrophils and undetectable for the hepatocytes. The binding of ^{32}P -inositol-1,4-bisphosphate (^{32}P -Ins(1,4) P_2) purified from the red cell hydrolysate, was also examined; total binding to neutrophils and hepatocytes was similar to that for ^{32}P -Ins(1,4,5) P_3 but binding was not inhibited by $10\text{ }\mu\text{M}$ of either Ins(1,4) P_2 or Ins(1,4,5) P_3 . Binding of ^{32}P -Ins(1,4,5) P_3 was not inhibited by $10\text{ }\mu\text{M}$ Ins(1,4) P_2 , inositol 4,5-bisphosphate (Ins(4,5) P_2), inositol-1-phosphate or cyclic inositol-1,2-phosphate. In hepatocytes, omission of ATP or reduction of bovine serum albumin from 2 to 0.2% did not affect specific binding. Increasing the Ca^{2+} concentration from 180 nM to $1.0\text{ }\mu\text{M}$ slightly increased specific binding ($125 \pm 10\%$ of control, $n=4$),

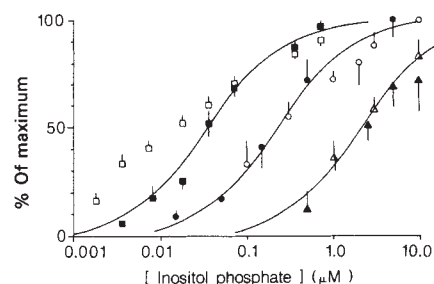


Fig. 2 Receptor occupancy (inhibition of ^{32}P -Ins(1,4,5) P_3 binding) and $^{45}\text{Ca}^{2+}$ -releasing activity of Ins(1,4,5) P_3 and Ins(2,4,5) P_3 in guinea pig hepatocytes, and of Ins(1,4,5) P_3 in rabbit neutrophils. Binding to permeable cells was determined as described for Fig. 1. Occupancy was calculated as per cent inhibition of specific ^{32}P -Ins(1,4,5) P_3 binding. Release of $^{45}\text{Ca}^{2+}$ for the rabbit neutrophils was determined as described previously¹⁰ and the data expressed as a percentage of maximal release ($50 \pm 5\%$ of a total ATP-dependent $^{45}\text{Ca}^{2+}$ uptake of $1.02 \pm 0.14\text{ nmol per mg protein}$). The data for release of ^{45}Ca from guinea pig hepatocytes are taken from ref. 10, and are expressed as per cent maximal release with Ins(1,4,5) P_3 . The curves are theoretical curves for a single-site interaction with K_D equal to the graphically estimated EC_{50} for ^{45}Ca release. Δ , $\blacktriangle$, Ins(2,4,5) P_3 in hepatocytes; $\circ$, $\bullet$, Ins(1,4,5) P_3 in hepatocytes; $\square$, $\blacksquare$, Ins(1,4,5) P_3 in neutrophils; $\triangle$, $\circ$, $\square$, occupancy; $\blacktriangle$, $\bullet$, $\blacksquare$, ^{45}Ca release. Values shown are mean $\pm$ s.e.m. of 3–6 observations each, except for the ^{45}Ca release from the neutrophils at 3.6 nM Ins(1,4,5) P_3 , which is the mean of two observations.

and reducing the Ca^{2+} concentration to 20 nM slightly reduced binding ($82 \pm 6\%$ of control, $n=6$). For hepatocytes, 100 and $300\text{ }\mu\text{M}$ Ins(4,5) P_2 inhibited binding by 45 and 85%, respectively, in two experiments.

Figure 2 shows the concentration dependence of inhibition of binding of ^{32}P -Ins(1,4,5) P_3 (indicated as per cent occupancy) by non-radioactive Ins(1,4,5) P_3 in permeable neutrophils and hepatocytes and by inositol-2,4,5-trisphosphate (Ins(2,4,5) P_3) in hepatocytes. Also shown are the concentration dependencies of $^{45}\text{Ca}^{2+}$ release by inositol trisphosphates under similar conditions. There was a close correlation between the estimated fractional receptor occupancies of Ins(1,4,5) P_3 and Ins(2,4,5) P_3 and ^{45}Ca release in the hepatocyte. The curves in Fig. 2 were generated by taking the observed EC_{50} (50% effective concentration) for ^{45}Ca release as a dissociation constant (K_D) for binding and by assuming a single binding site with a Hill coefficient of 1. Although variability in the data may not permit detection of subtle deviations from the single-site model, the fit appears adequate for the hepatocyte.

In neutrophils, Ins(1,4,5) P_3 was more potent in releasing Ca^{2+} and in displacing bound ^{32}P -Ins(1,4,5) P_3 than in hepatocytes. However, low concentrations of Ins(1,4,5) P_3 ($<30\text{ nM}$) could displace proportionately greater binding than predicted either from the $^{45}\text{Ca}^{2+}$ data or from the single-site binding model. This indicates either complex binding kinetics to a single site, or binding to more than one site. Considering the latter possibility, a high-affinity site may exist which is not coupled to Ca^{2+} release; we speculate that this site indicates some other regulatory function of Ins(1,4,5) P_3 in the neutrophil not found in the hepatocyte.

The relatively low affinity of Ins(1,4,5) P_3 for its receptor when compared with more conventional ligands used in binding assays requires high-specific-activity ^{32}P -Ins(1,4,5) P_3 to detect specific binding. Because of limited quantities of the radioligand and the necessary dilution of the specific radioactivity at higher concentrations, Scatchard analysis of the data was not appropriate. Receptor density was estimated for the hepatocyte (which fits the single-site model) by extrapolating the most reliable binding data (at 100 and 300 nM) with the following relation: $B_{\text{max}} = B(K_D + [S])/[S]$ (where B_{max} is maximum binding capacity, K_D is assumed to be equal to the EC_{50} for the

hepatocytes (220 nM; ref. 10) and B is the concentration of bound ligand at ligand concentration $[S]$. At 100 nM $\text{Ins}(1,4,5)\text{P}_3$, binding was 59 ± 8 ($n=6$) fmol per mg protein, and at 300 nM, binding was 115 ± 20 ($n=5$) fmol per mg protein. These values give receptor densities of 189 and 199 fmol per mg protein, respectively, with an average of 194 fmol per mg protein. As the guinea pig hepatocyte contains $2.18 \mu\text{l}$ intracellular water per mg protein¹³, the K_D for $\text{Ins}(1,4,5)\text{P}_3$ (220 nM) expressed relative to cellular protein in intact cells, would be 480 fmol per mg protein. That the estimated receptor density is less than the K_D of the receptor may be significant in that the receptor would not represent a significant depot of $\text{Ins}(1,4,5)\text{P}_3$ binding. The intracellular concentrations of free and bound $\text{Ins}(1,4,5)\text{P}_3$ and resulting Ca^{2+} mobilization would thus be approximately proportional to the rate of $\text{Ins}(1,4,5)\text{P}_3$ production and to receptor activation, at least at submaximal concentrations of agonists.

In summary, the available data indicate that a binding site for $\text{Ins}(1,4,5)\text{P}_3$ can be demonstrated using high-specific-activity $^{32}\text{P}\text{-Ins}(1,4,5)\text{P}_3$, and that this site is the physiological receptor for $\text{Ins}(1,4,5)\text{P}_3$. Binding is specific, saturable and reversible. Inositol phosphates that have Ca^{2+} -releasing activity¹⁰ ($\text{Ins}(1,4,5)\text{P}_3$, $\text{Ins}(2,4,5)\text{P}_3$ and $\text{Ins}(4,5)\text{P}_2$) bind to this site, whereas those which do not release Ca^{2+} (ref. 10) do not bind. Furthermore, with the exception of some very high-affinity binding in the neutrophil, there is a close correlation between the ability of $\text{Ins}(1,4,5)\text{P}_3$ and $\text{Ins}(2,4,5)\text{P}_3$ to bind to the receptor and the abilities of these molecules to release Ca^{2+} . Receptors for ligands on the surface of cells are known to be regulated in several ways¹⁴, and it is likely that such control mechanisms

also occur at an intracellular receptor. The demonstration and quantitation of an intracellular receptor for the Ca^{2+} -mobilizing messenger, $\text{Ins}(1,4,5)\text{P}_3$, should open the way for future investigations into the nature and regulation of this ubiquitous site in various biological systems.

We thank R. F. Irvine for non-radioactive inositol phosphates. This research was supported by NIH grants AM-28029 and AM-32823. A.S. was supported in part by a fellowship from the Educational Commission for foreign medical graduates.

Received 16 September; accepted 3 December 1985.

- Downes, P. & Michell, R. H. *Cell Calcium* **3**, 467-502 (1982).
- Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
- Berridge, M. J. *Biochem. J.* **212**, 849-858 (1983).
- Streb, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67-68 (1983).
- Burgess, G. M. *et al.* *Nature* **309**, 63-66 (1984).
- Suematsu, E., Hirata, M., Hashimoto, T. & Kuriyama, H. *Biochem. biophys. Res. Commun.* **120**, 481-485 (1984).
- Prentki, M. *et al.* *Nature* **309**, 562-564 (1984).
- Streb, H., Bayerdorffer, E., Haase, W., Irvine, R. F. & Schulz, I. *J. Membrane Biol.* **81**, 241-253 (1984).
- Dawson, A. P. & Irvine, R. F. *Biochem. biophys. Res. Commun.* **120**, 858-864 (1984).
- Burgess, G. M., Irvine, R. F., Berridge, M. J., McKinney, J. S. & Putney, J. W. Jr *Biochem. J.* **224**, 741-746 (1984).
- Downes, C. P. & Michell, R. H. *Biochem. J.* **198**, 133-140 (1981).
- Burgess, G. M., McKinney, J. S., Fabiato, A., Leslie, B. A. & Putney, J. W. Jr *J. Biol. Chem.* **258**, 15336-15345 (1983).
- De Witt, L. M. & Putney, J. W. Jr *J. Physiol., Lond.* **346**, 395-407 (1984).
- Czech, M. P. & Kahn, C. R. (eds) *Membrane Receptors and Cellular Regulation* (Liss, New York, 1985).
- Bradford, P. G. & Rubin, R. P. *Molec. Pharmac.* **27**, 74-78 (1985).
- Quist, E. *Archs Biochem. Biophys.* **236**, 140-149 (1985).
- Berridge, M. J., Dawson, R. M. C., Downes, C. P., Heslop, J. P. & Irvine, R. F. *Biochem. J.* **212**, 473-482 (1983).
- Desjournet, A. & Petek, F. *Bull. Soc. Chim. Biol.* **38**, 871-883 (1956).
- Irvine, R. F., Anggard, E. A., Letcher, A. J. & Downes, C. P. *Biochem. J.* **229**, 505-511 (1985).

Stimulation of bone resorption and inhibition of bone formation *in vitro* by human tumour necrosis factors

Donald R. Bertolini, Glenn E. Nedwin*†, Timothy S. Bringman*, Donna D. Smith & Gregory R. Mundy

Department of Medicine, Division of Endocrinology, University of Texas Health Science Center, San Antonio, Texas 78284, USA

* Genentech Inc., South San Francisco, California 94080, USA

When leukocytes are exposed to mitogens or antigens *in vitro*, they release bone-resorbing activity into the culture supernatants which can be detected by bioassay^{1,2}. Like many lymphocyte-monocyte products, this activity has been difficult to purify because of its low abundance in activated leukocyte cultures and the unwieldy bioassay required to detect biological activity. Partially purified preparations of this activity inhibit bone collagen synthesis in organ cultures of fetal rat calvariae³. Recent data suggest that both activated lymphocytes and monocytes release factors which could contribute to this activity². Recently, monocyte-derived tumour necrosis factor α (TNF- α) and lymphocyte-derived tumour necrosis factor β (TNF- β) (previously called lymphotoxin), two multifunctional cytokines which have similar cytotoxic effects on neoplastic cell lines, have been purified to homogeneity^{4,5} and their complementary DNAs cloned and expressed in *Escherichia coli*^{6,7}. As both of these cytokines are likely to be present in activated leukocyte supernatants, we tested purified recombinant preparations for their effects on bone resorption and bone collagen synthesis *in vitro*, and report here that both cytokines at 10^{-7} to 10^{-9} M caused osteoclastic bone resorption and inhibited bone collagen synthesis. These data suggest that at least part of the bone-resorbing activity present in activated leukocyte culture supernatants may be due to these cytokines.

Recombinant human TNF- α (rTNF- α)⁷ and human TNF- β (rTNF- β)⁶ used in this study were expressed in *E. coli* and the recombinant molecules purified as described previously^{6,7}. The rTNF- α and rTNF- β stimulated bone-resorbing activity when incubated with fetal rat bones in which ⁴⁵Ca had been incorporated previously using the technique described by Raisz⁸. This bone-resorbing activity was dose-dependent from 10^{-7} to 10^{-9} M (Fig. 1). Maximal activity for both cytokines was observed at 10^{-7} M (Fig. 1). The time course of action of the cytokines was examined by incubating the bones in the presence of control medium or medium with the appropriate cytokine for 2 days and then transferring the bones to fresh control medium or medium with the cytokine for a further 3 days (data not shown). The factors released approximately equal amounts of calcium over the two time periods, similar to that shown previously for the bone-resorbing activity in non-fractionated activated leukocyte culture supernatants and parathyroid hormone (PTH)^{3,9-11}.

Treatment of bone cultures with rTNF- α and rTNF- β caused an increase in the number of multinucleated osteoclasts and a decrease in the amount of mineralized bone matrix (Fig. 2). In contrast, few osteoclasts were present in control bones. As further evidence that rTNF- α and rTNF- β cause resorption by osteoclast activation, their effects on bone resorption at 10^{-7} M were completely inhibited by salmon calcitonin (5 units ml⁻¹), a specific inhibitor of osteoclast activity (Table 1).

The effects of recombinant human TNF- α and TNF- β on bone collagen synthesis were determined using the method of Peterkofsky and Diegelman¹², which measures the incorporation of ³H-proline into collagenase-digestible protein. The culture technique for studying factors which influence collagen synthesis *in vitro* has been described previously^{13,14}. We used the modified procedure described by Kream *et al.*¹⁴, which uses cultured frontal and parietal bone explants from 21-day fetal rats. Both rTNF- α and rTNF- β inhibited collagenase-digestible protein synthesis to a greater extent than they inhibited non-collagen protein synthesis (Table 1). Thus, the per cent collagen synthesis was significantly inhibited compared with corresponding bones incubated with control media.

† Present address: Molecular Therapeutics, Inc., 400 Morgan Lane, West Haven, Connecticut 06516, USA.

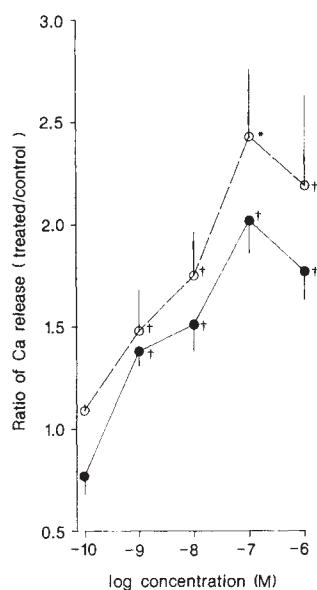


Fig. 1 Effects of recombinant TNF- α and TNF- β on bone resorption *in vitro*. rTNF- α (○) or rTNF- β (●) was diluted into fresh BGJ medium with 5% fetal bovine serum (FBS) and incubated with fetal rat long bones for 48 h. Data shown are the ratio of medium ^{45}Ca released by bones treated with test media to that released by bones treated with corresponding control media; each value represents the mean $\pm$ s.e.m. for four pairs of bone cultures. PTH (200 ng ml $^{-1}$) was used as a positive control in this experiment and gave a ratio of 2.33 ± 0.70 . Statistical differences from 1.0 at each concentration were determined using Student's *t*-test for non-paired samples (molarity calculation based on monomer M_r).

* $P < 0.005$; † $P < 0.05$.

We also tested the effects of recombinant human cytokines on alkaline phosphatase activity in cells with the osteoblast phenotype (17/2.8 rat osteosarcoma cells) according to the method of Majeska *et al.*¹⁵. Alkaline phosphatase is a marker enzyme of new bone formation; agents which inhibit bone formation cause a decrease in the alkaline phosphatase content of cells of the osteoblast phenotype. The cells were incubated with the test substances for 36–72 h, and the enzyme content was assayed according to the method of Lowry *et al.*¹⁶. Both cytokines, in the concentration range 10^{-6} – 10^{-10} M, caused a decrease in the alkaline phosphatase content of cells with the osteoblast phenotype. Enzyme content was inhibited 50% by both cytokines at 10^{-10} M, and 85% by both cytokines at 10^{-7} M. In the same experiment, PTH (50 ng ml $^{-1}$) also inhibited enzyme content by 85%.

The bone-resorbing activity present in activated leukocyte culture supernatants has been referred to as osteoclast activating factor (OAF), but appears to represent a family of proteins from different cellular sources. OAF was originally thought to be a lymphocyte product that required the presence of monocytes in the culture for its production^{1,17} and it is possible that TNF- β is the bone-resorbing activity that was produced in these studies using separated cell populations or lymphoid cell lines. However, monocyte products such as interleukin-1 (ref. 18) and prostaglandins¹⁹ have been shown to stimulate osteoclastic bone resorption *in vitro*²⁰, and these factors presumably contribute to the bone-resorbing activity present in activated leukocyte culture supernatants.

Malignant diseases of lymphoid cells such as myeloma, T-cell lymphoma and Burkitt's lymphoma are frequently associated with bone destruction and sometimes with hypercalcaemia²¹. Freshly isolated cells and cultured cell lines, such as the myeloma cell line RPMI 8226, derived from patients with these disorders, release bone-resorbing activity having similar chemical and bio-

Table 1 Effects of salmon calcitonin (CT) on bone resorption stimulated by human recombinant tumour necrosis factors

Treatment of bone cultures	Bone-resorbing activity (per cent ^{45}Ca release)
Control media	22 $\pm$ 2
Control media + CT (5 U ml $^{-1}$)	13 $\pm$ 1*
TNF- α (10^{-7} M)	38 $\pm$ 3
TNF- α (10^{-7} M) + CT (5 U ml $^{-1}$)	11 $\pm$ 1†
TNF- β (10^{-7} M)	33 $\pm$ 3
TNF- β (10^{-7} M) + CT (5 U ml $^{-1}$)	12 $\pm$ 1†

The bones were cultured for 48 h. Calcitonin was added together with the cytokine at the start of the culture period. Values are means $\pm$ s.e.m. for four bone cultures.

*† Significantly less ($P < 0.025$ and < 0.005 , respectively) than corresponding bone cultures not incubated with CT.

Table 2 Effects of rTNF- α and rTNF- β on ^3H -proline incorporation into collagenase-digestible protein (CDP) and non-collagen protein (NCP) in fetal rat calvaria

Treatment		CDP (d.p.m. per μg dry wt)	NCP (d.p.m. per μg dry wt)	% Collagen synthesis*
Control		86 $\pm$ 7	171 $\pm$ 12	8.5 $\pm$ 0.3
TNF- α	10^{-7} M	19 $\pm$ 2	88 $\pm$ 7	3.9 $\pm$ 0.3†
	10^{-8} M	32 $\pm$ 6	114 $\pm$ 13	5.0 $\pm$ 0.6†
	10^{-9} M	44 $\pm$ 5	147 $\pm$ 15	5.3 $\pm$ 0.3†
TNF- β	10^{-7} M	37 $\pm$ 3	128 $\pm$ 4	5.0 $\pm$ 0.3†
	10^{-8} M	33 $\pm$ 5	116 $\pm$ 8	4.9 $\pm$ 0.6†
	10^{-9} M	51 $\pm$ 6	133 $\pm$ 8	6.6 $\pm$ 0.7‡
Insulin	10^{-6} M	180 $\pm$ 23	205 $\pm$ 17	13.9 $\pm$ 1.0†

Values are mean $\pm$ s.e.m. for six half-calvariae incubated for 48 h in the conditions indicated and pulsed with ^3H -proline for the final 4 h.

* Proportion of proline incorporated into collagen is corrected for the relative abundance of proline in CDP and NCP.

†‡ Significantly different from control ($P < 0.005$ and < 0.05 , respectively).

logical characteristics to the bone-resorbing activity present in activated leukocyte culture supernatants^{3,9,22}. Myeloma cell lines, including 8226, have also been shown to produce TNF- β ²³. Osteoclast activation in both organ cultures incubated with rTNF- α and rTNF- β resembles what is found in the bones of patients with myeloma. Moreover, it is characteristic of these malignancies that there is no evidence of new bone formation associated with the increased bone resorption, in contrast to what is usually seen in osteolytic bone lesions associated with solid tumours or with primary hyperparathyroidism¹⁷. Whether either or both of these cytokines are the only bone-resorbing factors released by malignant lymphoid cells associated with hypercalcaemia will require further study.

TNF- α and TNF- β have biochemical characteristics similar to those of the bone-resorbing activity present in activated leukocyte culture supernatants. Like TNF- α and TNF- β , OAF is produced by activated leukocytes, is heat-sensitive and is inactivated by protease²⁴. OAF has a relative molecular mass (M_r) of 14,000 (14K)–25,000 (25K), estimated by gel permeation chromatography²⁴, but under certain conditions it can run with a M_r of between 45K and 70K^{3,24}. Purified TNF- α has a M_r of 17.8K and is 157 amino acids long but elutes from gel permeation columns as a 45K protein. Purified TNF- β has a M_r of 18.7K and is 171 amino acids long but elutes from gel permeation columns as a 60K protein⁵. Natural TNF- β is a glycoprotein which is probably glycosylated at asparagine residue 62 (ref. 5). As TNF- β expressed in *E. coli* is not glycosylated, our data indicate that glycosylation is not necessary for bone-resorbing activity. It has also been demonstrated that glycosylation is not necessary for the cytotoxic effect of TNF- β on tumour cells⁶.

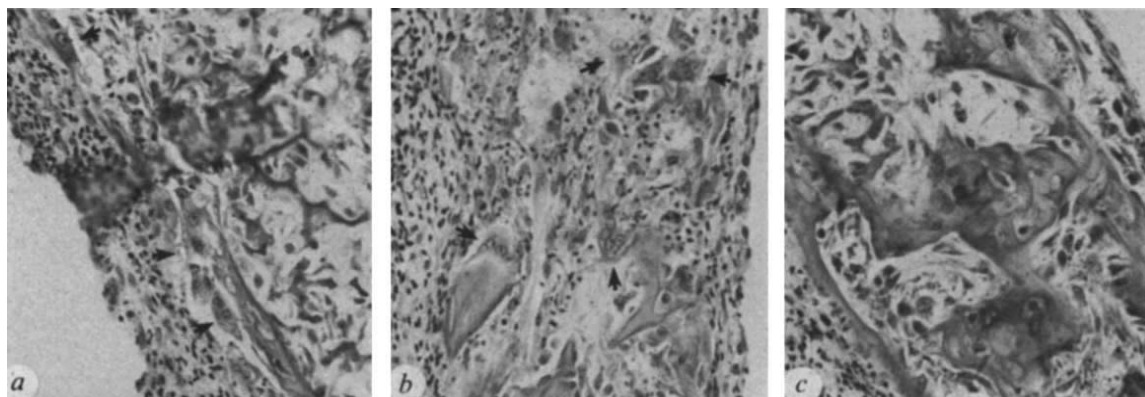


Fig. 2 Histological sections of fetal rat long bones cultured for 48 h with $\text{TNF-}\alpha$ (a), $\text{TNF-}\beta$ (b) or control medium (c), showing increased osteoclast number and activity and decreased mineralized matrix in bones treated with the cytokines. Some of the obvious osteoclasts are indicated by arrows. The cytokines were incubated with the bone cultures at a concentration of 10^{-7} M.

As $\text{TNF-}\alpha$ and $\text{TNF-}\beta$ are present in activated leukocyte cultures and both are potent bone resorbers, we suggest that they may account for a part (or all) of the bone-resorbing activity presently referred to as OAF.

Testing of rTNFs *in vitro* may reveal that they have a number of other biological activities which could be important in patients with malignant disease. Recently, it has been shown that $\text{TNF-}\alpha$ may be identical to cachectin²⁵, the macrophage product which inhibits the activity of lipoprotein lipase in adipocytes²⁶, and which may be responsible for the wasting that occurs in patients with malignant disease. Moreover, $\text{TNF-}\alpha$ may also mediate the tissue effects of the shock syndrome caused by endotoxin²⁷, a lipopolysaccharide component of the cell walls of some bacteria. It is imperative that any other biological activities of these molecules are identified as they are now being used in clinical trials in patients with malignant disease.

The present results may have important implications for the control of normal bone remodelling. The cellular sequence of events responsible for bone remodelling begins with a local increase in osteoclast activity. The factors responsible for this local osteoclast activation are unknown, but since activation of bone remodelling units occurs in discrete areas, it seems likely that local factors in the bone marrow microenvironment are responsible. Osteoclast activating factor has been proposed as a local mediator of osteoclast activity. The identification of specific bone-resorbing lymphokines and monokines for which cDNA probes are available may allow us to determine the role of these factors in normal bone remodelling and the increase in osteoclast activity which occurs in age-related bone loss.

We thank Genentech, Inc. for providing rTNFs, Dr Maxine Gowen for helpful criticisms and Nancy Garrett for preparing the manuscript. This work was supported by NIH grants CA29537, AM28149, AM07464 and CA40035.

Note added in proof: Since submission of this manuscript, Dewhirst *et al.*²⁸ have reported purification of interleukin-1 from short-term phytohaemagglutinin-activated leukocyte cultures using a bone resorption assay. However, the relative roles of

interleukin-1 and the TNFs as OAFs may depend on conditions such as duration of the leukocyte culture, antigen or mitogen used as stimulant, and types of leukocytes present in the culture.

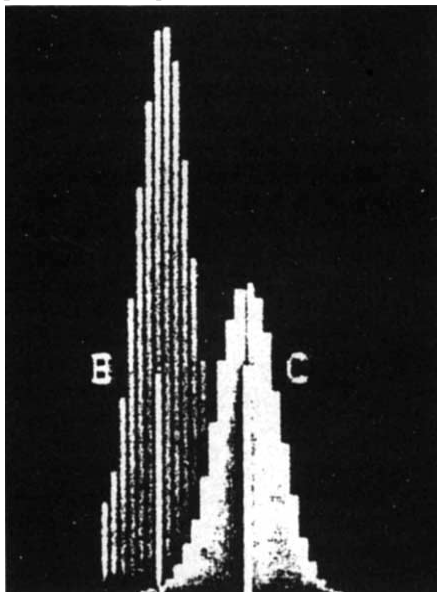
Received 30 August; accepted 2 December 1985.

- Horton, J. E., Oppenheim, J. J., Mergenhagen, S. E. & Raisz, L. G. *J. Immun.* **113**, 1278-1287 (1974).
- Mundy, G. R. & Bertolini, D. R. in *Endocrinology, int. Congr. Ser.* 665 (eds Labrie, F. & Proulx, L.) 587-590 (Excerpta Medica, Amsterdam, 1984).
- Raisz, L. G. *et al. J. clin. Invest.* **56**, 408-413 (1975).
- Aggarwal, B. B., Henzel, W. J., Moffat, B., Kohr, W. J. & Harkins, R. N. *J. biol. Chem.* **260**, 2334-2344 (1985).
- Aggarwal, B. B. *et al. J. biol. Chem.* **260**, 2345-2354 (1985).
- Gray, P. W. *et al. Nature* **312**, 721-724 (1984).
- Pennica, D. *et al. Nature* **312**, 724-729 (1984).
- Raisz, L. G. *J. clin. Invest.* **44**, 103-116 (1965).
- Mundy, G. R., Luben, R. A., Raisz, L. G., Oppenheim, J. J. & Buell, D. N. *New Engl. J. Med.* **290**, 867-871 (1974).
- Mundy, G. R. & Raisz, L. G. *J. clin. Invest.* **60**, 122-128 (1977).
- Horton, J. E., Raisz, L. G., Simmons, H. A., Oppenheim, J. J. & Mergenhagen, S. E. *Science* **177**, 793-795 (1972).
- Peterkofsky, B. & Diegelmann, R. F. *Biochemistry* **10**, 988-994 (1971).
- Dietrich, J. W., Canalis, E. M., Maina, D. M. & Raisz, L. G. *Endocrinology* **98**, 943-959 (1976).
- Kream, B. E., Smith, M. D., Canalis, E. & Raisz, L. G. *Endocrinology* **116**, 296-302 (1985).
- Majeska, R. J., Rodan, S. B. & Rodan, G. A. *Expl Cell Res.* **111**, 465-468 (1978).
- Lowry, O. H. *Meth. Enzym.* **4**, 371-372 (1955).
- Chen, P., Trummel, C., Horton, J., Baker, J. J. & Oppenheim, J. J. *Eur. J. Immun.* **6**, 732-736 (1976).
- Gowen, M., Wood, D. D., Ihrle, E. J., McGuire, M. K. B. & Russell, R. G. G. *Nature* **306**, 378-380 (1983).
- Dominguez, J. H. & Mundy, G. R. *Calcif. Tissue Int.* **31**, 29-34 (1980).
- Klein, D. C. & Raisz, L. G. *Endocrinology* **86**, 1436-1440 (1970).
- Mundy, G. R. *et al. New Engl. J. Med.* **310**, 1718-1727 (1984).
- Mundy, G. R., Raisz, L. G., Cooper, R. A., Schechter, G. P. & Salmon, S. E. *New Engl. J. Med.* **291**, 1041-1046 (1974).
- Nedwin, G. E., Svedersky, T. S., Bringman, T. S., Palladino, M. A. & Goeddel, D. V. *J. Immun.* **135**, 2492-2497 (1985).
- Luben, R. A., Mundy, G. R., Trummel, C. L. & Raisz, L. G. *J. clin. Invest.* **53**, 1473-1480 (1974).
- Beutler, B. *et al. Nature* **316**, 552-554 (1985).
- Torti, F. M., Dieckmann, B., Beutler, B., Cerami, A. & Ringold, G. M. *Science* **229**, 867-869 (1985).
- Beutler, B., Milsark, I. W. & Cerami, A. C. *Science* **229**, 869-871 (1985).
- Dewhirst, F. E., Stasheuko, P. P., Mole, J. E. & Tsurumachi, T. *J. Immun.* **135**, 2562-2568 (1985).

Microanalysis gets new pulse processor

Heading this month's selection of lasers, software and new literature is an X-ray processor that can detect elements as light as boron.

KEVEX Corporation's model 4460 pulse processor (Reader Service No. 100) promises to erase the noise peaks that plagued X-ray microanalysis, giving **light element sensitivity sufficient for boron detection**. Improved signal-to-noise ratios allow all low-energy peaks, especially the carbon peak, to be more clearly resolved. Kevex improved upon conventional gaussian electronics by using triangular pulse-shaping circuitry instead. The resulting processor handles a higher X-ray throughput for better precision in less time. The



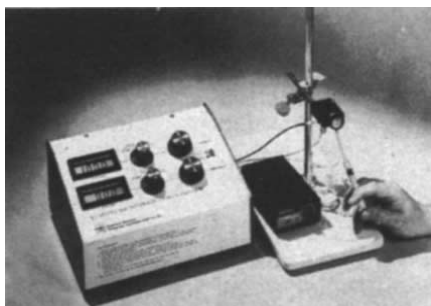
The Kevex processor: new resolution for carbon and boron

Kevex model 4460, designed for the model 8000 microanalysis system, sells for \$2,200 in the United States. The 4560 was developed for model 7000 analysers.

Metres and monitors

Monitoring oxygen uptake or evolution in samples with low substrate volumes is the key feature of the biological oxygen monitor from Yellow Springs Instrument Co. Inc. (Reader Service No. 101). The YSI 5300, fitted with micro or standard oxygen probes, can run 600 μ l to 50 ml samples in batch or flow-through modes. The instrument can process two experiments simultaneously using two channels, displays and recorder outputs. This function allows the user to plot the differential of the two channels. YSI prices the 5300

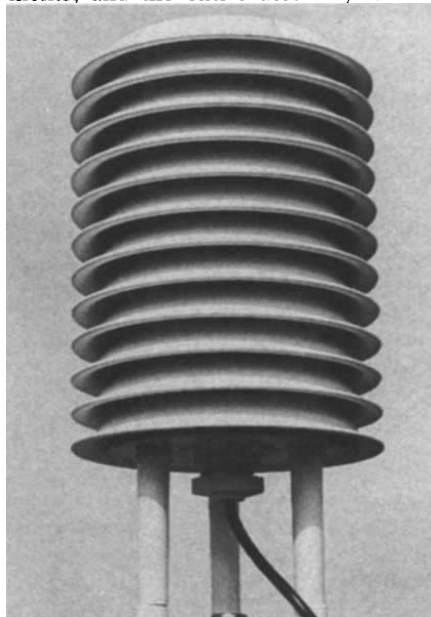
These notes are based on information provided by the manufacturers. To obtain further details about these products use the reader service card bound inside the journal.



Keeping tabs on oxygen with YSI's monitor

monitor and accessory package at about \$1,600.

For **relative humidity and temperature measurements** in remote and rugged environments, Hollis Geosystems Corp. has developed the durable RHT-241 (Reader Service No. 102). Originally intended for meteorological, agricultural or marine interests, the RHT-241's design incorporates low power consumption with rapid response time. A self-aspirating solar radiation shield protects the sensing elements, and the entire assembly is non-



Reliable readings in hostile territory from Hollis Geosystems

corrosive. The unit will survive dry, dusty climates or marine environments for at least one year. The RHT-241 measures 12 cm in diameter and 15 cm in height and costs about \$1,200.

A new portable concentration meter from Orion Research Inc. makes **measurement by ion-selective electrode**

(ISE) as easy as pH determination (Reader Service No. 103). The lightweight SA 270, marketed for \$650 in the United States, makes it possible to perform rapid analysis on-site. The unit prompts the user through a one- or two-point standardization, and the display reads in p.p.m., molarity, per cent and so on. Readout can



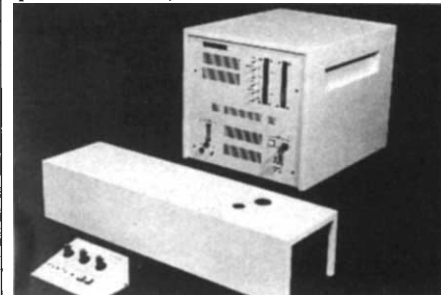
Orion's concentration metre features automatic temperature control

be in two or three digits, depending on the speed and accuracy desired.

The light fantastic

The latest addition to Laser Science Inc.'s compact sealed nitrogen and dye lasers is the VSL-337ND **near-diffraction nitrogen laser** (Reader Service No. 104). The laser's average power (after 10 million pulses) is 3.0 mW; its peak power is 40 kW and its pulse energy greater than 150 μ J. The 4x4 mm beam is focusable to less than 5 μ m, with a beam divergence of less than 0.3 mrad. The laser measures about 14x8x3 inches and operates at 337 nm. In the US, the 18-pound VSL-337ND can be had for under \$5,300.

The Omnicrome 500X **air-cooled argon laser** from Lambda Photometrics Ltd. is equipped with a hard-sealed metal-ceramic plasma tube with a life expectancy of greater than 10,000 hours, and a cast resonator that can withstand three axis shock loads exceeding 30 g for 10 milliseconds with no detectable change in performance (Reader Service No. 105).



The Omnicrome laser can be refurbished for prolonged lifetime

Priced at £4,911, the Omnicrome 500X is compatible with most other air-cooled argon lasers and the laser head is completely interchangeable with the Spectra-Physics 162 and ALC60X.

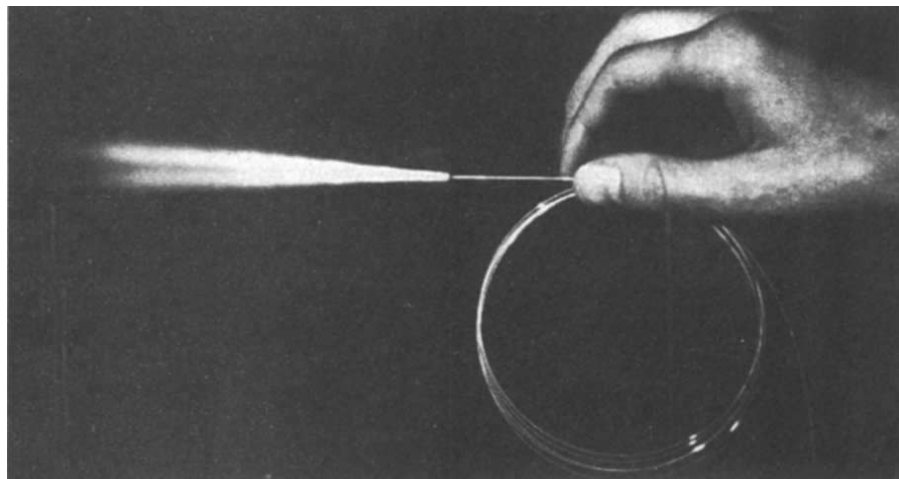
The Spectra-Physics DCR-11 Nd:YAG pulsed laser is ideal for portable applications and takes up much less space in the lab (Reader Service No. 106). The laser head is 33×9×10 inches and the power supply dimensions are 23×19×21 inches. At 1,064 nm, the DCR-11 has an output of 275 mJ and a repetition rate of 10 pulses per second. The output is nearly diffraction limited for efficient wavelength conversion: 135 mJ at 532 nm, 60 mJ at 355 nm, 30 mJ at 266 nm with the optional



Spectra-Physics's pulsed laser is one-third the size of previous lasers

harmonic generator. The laser sells for \$25,500 in the US.

Conventional lasers must be optically straight, rigid and have accurately aligned mirrors at each end, but the Optical Fibre



The rare-earth fibre laser is a new breed in optical electronics

Research Group at Southampton University has produced **flexible lasers**, thin as human hair and carried through the core of optical fibres (Reader Service No. 107). Normally, rare earth elements in the fibre core cause large optical losses, but the group found transparent windows where the transmission is just as good as in conventional telecommunications fibres. With the particular advantages of fibre lasers, namely axial pumping, high light intensity in the core and long lengths, it is now possible to obtain laser action with new materials.

The terminal enthusiast

The Continuous Flow Analyzer software from Spectra-Physics **quantitates the peaks obtained from the continuous flow analysis** of up to 120 samples at a time (Reader Service No. 108). This Autolab BASIC program requires a detector signal at regular intervals throughout a run and comes on a PROM chip for the SP4200 computing integrator. At \$2,000 in the United States, the software features automatic drift correction, peak overlap correction, statistics calculations, use of both peak areas and heights and peak assignment modification.

ELISA data analysis need no longer be done by hand. Bio-Rad Laboratories's MacReader software **eliminates the manual calculations involved in ELISA screen-**



McReader software comes in several languages ing and quantitation, and generates printed reports such as standard curve graphics, sample extrapolations and complete statistical analyses of data (Reader Service No. 109). MacReader links the

model 2550 automatic enzyme immunoassay reader with the Apple Macintosh computer and retails for \$600 in the US.

Unnecessary HPLC experiments may be eliminated with Drylab I, a computer program from LC Resources Inc. that simulates chromatographic runs (Reader Service No. 110). Based on the work of Stout, Snyder *et al.* this software uses the data from an initial separation to determine the optimum parameters for further separations (column plate number, resolution, etc). DryLab I can also help troubleshoot separation problems. For \$450, the software package consists of one hour of video instruction, a 100-page user's manual and the software itself.

Thermal analysis software is available from Mettler Instruments AG (Reader Service No. 111). The Mettler software combines the graphics capabilities of an IBM personal computer with the analytical precision of the TA3000 thermal analysis unit. The package costs £2,100 in the UK.

The LABS database compiles detailed **information on over 30,000 laboratories** in the United States and Canada (Reader Service No. 112). LABS records the types of research and the names of personnel at each station so that researchers can identify groups with a combination of characteristics, such as location, size and area of interest. Created by Coalesce Corp., the database encompasses scientific, engineering and medical labs. Coalesce charges \$100 to \$300 for initial queries.

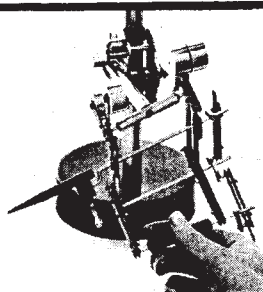
Paperbound expertise

Kratos Analytical offers a wall chart detailing the **basic principles of mass spectrometry** and the functioning of a mass spectrometer (Reader Service No. 113).

The new **fast-affinity chromatography (FAC)** technique is the subject of a 12-page publication from Beckman Instru-

ADVERTISEMENTS

Ideal for microinjection of xenopus oocytes



SINGER MICROMANIPULATORS ARE FAMOUS BECAUSE

Singer Micromanipulators enable anyone to make exact movements without tremor or lost motion, in three dimensions, using a single control handle under low or high power magnification.

A varied range available. All simple to operate.

SINGER INSTRUMENT CO. LTD.,
Treborough Lodge, Roadwater, Somerset.
Tel: Washford (0984) 40226/325

Reader Service No.33

PHORBOL ESTERS

99 + % PURE
PMA (TPA)

PHORBOL 12-MYRISTATE
13-ACETATE

Cat. #P1680 5mg/\$28⁰⁰



LC SERVICES CORPORATION

165 New Boston Street
Woburn, MA 01801
(617) 938-1700

Reader Service No.34

ments Inc. (Reader Service No. 114). Bulletin 5933 describes how to plan an FAC, with illustrations and step-by-step instructions for ligand selection, derivatization, sample loading and elution. The brochure highlights six practical examples using Beckman's Ultrafinity-EP column.

Gas, liquid or thin-layer chromatographers may want to take a look at Analytichem's **sample preparation handbook** (Reader Service No. 115). The 124-page book offers solutions to preparation problems, explains the technique of



Analytichem offers solutions to preparation problems

sorbent extraction, and spells out the thought processes that lead to successful development of a sample preparation method.

Lecture notes on the **basic principles of recombinant DNA** technology and HPLC applications in nucleotide analysis have been introduced by Varian Associates Ltd. (Reader Service No. 116). The 42-page booklet is aimed at HPLC users and lecturers who may find the notes useful in introducing students to nucleic acid and nucleotide chromatography.

Safety firsts

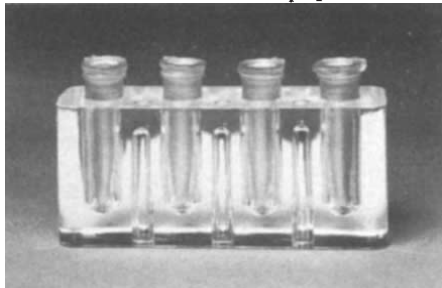
For the occasional laboratory mishap, Wheaton Instruments provides a **polycarbonate safety shield** to guard against spilled chemicals and broken apparatus (Reader Service No. 117). The 1/4-inch thick shield clamps to the bench top with an aluminium base. Wheaton carries two sizes, 20 1/4 x 18 inches and 26 3/8 x 20



A safety shield puts accidents at a distance

inches, for \$92 and \$144 respectively in the United States.

The acrylic Beta Blok from Research Products International Corp. **protects fingers and hands** from overexposure to beta radiation (Reader Service No. 118). The blocks come in cases of six, each block holding four 1.5 ml or 0.25 ml microcentrifuge tubes. Suitable for work with ³²P, the Beta Bloks sell for \$57 per case in the United States.

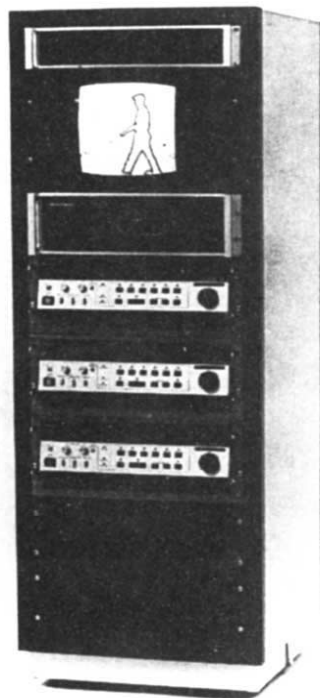


The Beta Blok makes the least of radiation

gers and hands from overexposure to beta radiation (Reader Service No. 118). The blocks come in cases of six, each block holding four 1.5 ml or 0.25 ml microcentrifuge tubes. Suitable for work with ³²P, the Beta Bloks sell for \$57 per case in the United States.

Science in motion

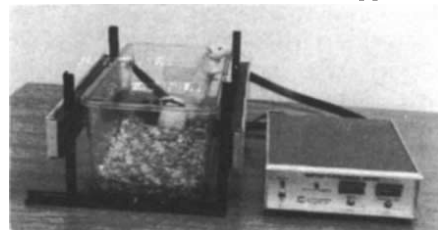
ExpertVision, a fully integrated computer system, **analyses and defines biochemical motion** with greater accuracy than standard film and hand digitizing methods (Reader Service No. 119). The system,



ExpertVision captures every step

designed by Motion Analysis Corp., merges computer and video technologies to enable two- or three-dimensional analysis on a high-speed graphics workstation or a personal computer. Offered at prices from \$30,000 to \$150,000 in the US, the system provides several advantages such as virtual elimination of marker 'drop-out', variable camera placement and programmable software.

Columbus Instruments markets a somewhat less complex **MINI animal activity-meter** for recording movements of creatures as small as insects or as large as rabbits (Reader Service No. 120). Fifteen infrared beams span the animal cage, with adjustable receivers and emitters to accommodate a variety of cage sizes. The MINI model B separates ambulatory from stereotypic activity and uses two electronic counters to display ambulatory and total animal activity. The MINI model B sells for \$1,995 in the United States and will interface with IBM-PC/XT or Apple IIe



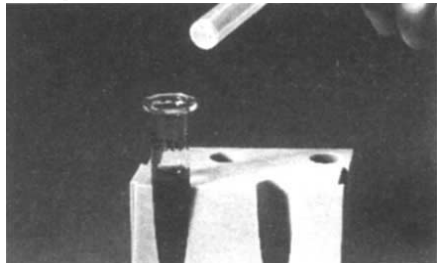
Animal activity surveillance from Columbus

computers through software also available from Columbus Instruments.

Fresh ideas, faithful methods

Separations and extractions previously performed with separatory funnels can now be carried out with a **new separation system** from Cole-Parmer Instrument Co. The Mixxor device effects separation via a mixer piston fitted into a graduated mixing reservoir and connected to a stoppered collection container (Reader Service No. 121). With five strokes of the piston the sample separates, and the lower phase stays in the mixer reservoir as the upper phase goes to the stoppered collecting container. The piston becomes the divider between the upper and lower phases. Mixxor sets come in five different capacities from 2 to 50 ml, and prices range from \$89 to \$139 in the United States.

A novel technique for making **unilamellar liposomes** has been advanced by Xydex Technologies Inc. (Reader Service No. 122). The glass Liposore cylinder creates lipid bilayers 0.1 to 1.0 µm in size in three to five minutes by forcing a phospholipid solution through capillary channels in a glass piston. The device is available in 2, 5



Liposomes made easy

and 10 ml sizes, for prices ranging from \$99 to \$119 in the United States. □

Breaks in career — What chance for women?

from Richard Pearson

Will demographic trends and skill shortages be more effective, at least for women in professional and managerial careers, than the 1960s and 1970s legislation in support of equal opportunities?

THE proportion of women working or seeking work is now over 48%, and until the recession of the early 1980s had been steadily growing from a figure of 1 in 10 in 1921. Women now account for 43% of the labour force in the United Kingdom. While many are concentrated in part-time and low-paid jobs, they form significant and growing proportions of those in teaching, medicine, science and personnel work. They also represent 45% of those entering higher education. Technological and structural change in the economy have brought about a decline in traditionally male-dominated heavy industry and an expansion in new industrial sectors, such as electronics and services where women are more likely to be found. Likewise, the growth in managerial and highly skilled jobs, and professions such as computing is enhancing women's job prospects, particularly those qualified at graduate level. Demographic trends over the next decade are also likely to lead to an increased participation of women in professional jobs. With the number of 18-year-olds declining by one-third over the next decade, employers are likely to turn increasingly to women in their search for

promoting the need for career breaks. A report by the Institute of Manpower Studies¹ shows, however, that while employers often talk of the costs of policies to help women combine work with a family, very few have measured the costs and lost investment in letting them go.

Although many employers exceed the statutory provision in many areas of employment legislation, they do not tend to do so for maternity leave. As such, the legislative standard tends to be the norm, and most employers and employees seemed to be clear about the provisions of statutory maternity leave, which can last up to 40 weeks. However, individual companies' administrative procedures were less clear, and often poorly handled, with the presumption that the career professional or managerial women should be able to find out everything for themselves.

Data on maternity leave and women leaving from career breaks are also not easy to find, employers often recording women taking normal maternity leave as leavers when they were still technically employees. As a result, some did not know how many were on leave at any one time, which makes monitoring difficult and can inhibit the planning of the women's return to work. The available evidence about the numbers returning suggested that only a small proportion did so, one of the highest being in a computer company where 30% returned to full-time and 15% to part-time work.

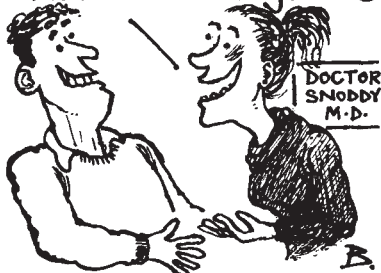
One major problem in managing the career break in the uncertainty as to the intention to return. There is a legal requirement for a woman to notify her employer of her intention to return, and to reconfirm it 7 weeks after the birth. Most automatically do this to safeguard their job, yet it can inhibit an employer's planning to cover the job and the timing of the subsequent return. One enlightened company has parallel discussions about the real intentions of each individual, without affecting the legal rights, a situation that requires a high degree of trust between the parties. A defeatist attitude to the legislation is clearly avoidable. Given an uncertain duration of leave, various strategies are used to cover the vacant job, including sharing the work out, hiring temps, or in larger organizations likely to have a number of similar jobs, filling it permanently. The latter can, however, cause problems on re-entry if there is not a similar job

available at the appropriate time, especially given the generally low levels of managerial turnover in most organizations. Where a woman is not returning to her original job, good practice should include early discussions and counselling as to the options; some women do not know what job they will do until they arrive, adding unnecessary anxiety to an already stressful period.

More women would return if it could be to part-time work, either as a transitional or as a more permanent arrangement. Some employers are able to accommodate this, and there are several computer software companies who operate with high percentages of part-time and often home-based, professional staff. In medicine, doctors have a national scheme which sets a quota of part-time jobs at senior registrar level, and also a retainer scheme which keeps doctors in touch with work if they wish to do very limited hours each week. In traditional organizations, attitudes of peers can be a problem. Thus, the case of a female doctor at consultant level where, in her speciality, work hours could be very flexible. It was quite acceptable for her male colleagues to organize work to allow time to play golf; it was not, however, seen as acceptable that she should organize her hours so as to collect her child from school!

To date, well-organized career breaks for women, let alone men, are still a rarity, yet they represent a missed opportunity for employers and individuals, particularly in the case of highly trained and experienced people. Barriers to change include a low realization of the costs of female wastage, and a failure to plan careers against a longer-term perspective. Resistance to women in management is still strong in many organizations, with a reluctance to accommodate flexible careers and part-time working in professional and managerial jobs. Pressure for a change from women is, however, growing, if slowly, and the more they move into such jobs and become a key source of skilled staff, the more labour market factors and economics will help reduce barriers to change. □

Darling—we're going to have a Maternity Leave!



talent. At the same time the birth bulge of the 1950s and '60s will see increasing numbers of 30–40-year-old women returning to the labour force in the 1990s.

How then are employers responding to these changes, and in particular what are they doing to help and even encourage women to return to their careers after a period of child-rearing? Those not responding positively may find themselves unable to retain an increasingly important component of their skilled workforce. At a minimum they will be letting go a huge investment in the training and development of employees with valuable and relevant skills.

The Engineering Council are strongly

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.

1. Hirsh, W., Hutt, R. & Atkinson, J. *Women, Career Breaks and Re-entry* (Institute of Manpower Studies, Brighton, 1985).